

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number
WO 02/20814 A1

(51) International Patent Classification⁷: **C12N 15/86**,
15/861, 5/10, 15/11, 15/63, 15/64, 15/65, A61K 48/00

Eun-Chung [KR/US]; 91 Waltham Street #1, Boston, MA 02118 (US). **WANG, Xinzhong** [CN/US]; 4 Woodmere Road, Framingham, MA 01701 (US).

(21) International Application Number: PCT/US01/27682

(74) Agent: **ELBING, Karen, L.**; Clark & Elbing LLP, 176 Federal Street, Boston, MA 02110-2214 (US).

(22) International Filing Date:
7 September 2001 (07.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/231,053 8 September 2000 (08.09.2000) US
60/246,904 8 November 2000 (08.11.2000) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 60/231,053 (CON)
Filed on 8 September 2000 (08.09.2000)
US 60/246,904 (CON)
Filed on 8 November 2000 (08.11.2000)

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **THE GENERAL HOSPITAL CORPORATION** [US/US]; 55 Fruit Street, Boston, MA 02214 (US).

Published:
— with international search report
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
— with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **SEED, Brian** [US/US]; 9 Hawthorne Place #5J, Boston, MA 02114 (US). **WRIGHT FREEMAN, Mason** [US/US]; 203 Lincoln Road, Lincoln, MA 01773 (US). **KOVTUN, Alexander** [US/US]; 156 Arlington Street, Acton, MA 01720 (US). **MURAKAWA, Masahiro** [JP/US]; 506 Beacon Street, Apt. 6, Boston, MA 02115 (US). **PARK,**

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SELF-REARRANGING DNA VECTORS

(57) Abstract: Disclosed are replicatable viral DNA vectors encoding a site-specific DNA-altering enzyme and a DNA target recognized by the enzyme, the enzyme selectively converting, in a cell expressing the enzyme, the DNA vector to a rearranged form. The invention further relates to methods for assembling recombinant adenoviral DNAs. These methods include the steps of: providing a first linearized DNA vector including a restriction site and a cos site and a second linearized DNA vector including the restriction site, an adenoviral nucleic molecule, and a cos site; and ligating the first and second linearized DNA vectors, the ligation assembling a recombinant adenoviral DNA.



WO 02/20814 A1

5

SELF-REARRANGING DNA VECTORS

Background of the Invention

The invention relates to DNA vectors.

10 Mammalian cell expression vectors based on DNA viruses have been widely discussed as gene delivery vehicles for genetic therapy. Among the different DNA viruses proposed for this purpose have been adenoviruses, baculovirus, Epstein Barr virus, and herpes simplex virus. In addition other smaller viruses that have an intranuclear phase in which the viral genome is present as a double stranded DNA, such
15 as retroviruses and parvoviruses, have been proposed as gene delivery vehicles.

Adenoviral vectors (AdV), for example, have a recognized potential for gene delivery, founded in their broad host range, robust growth in culture, and capacity to infect mitotically quiescent cells (Graham and Prevec, Manipulation of adenovirus vectors, p. 109-128, *In* E. J. Murray (ed.), Methods in Molecular Biology, vol. 7, Humana, Clifton, NJ, 1991; Trapnell and Gorziglia, Curr. Opin. Biotechnol. 5:617-625, 1994). AdV can be propagated in a helper cell line, 293, a human embryonic kidney cell line transformed by adenovirus type 5 (Graham et al., J. Gen.Virol. 36:59-72, 1994). 293 cells express the viral E1 gene products (E1a and E1b) that are the master regulatory proteins for subsequent viral gene expression. E1 deleted viruses can propagate in 293
20 cells, but not in other cells. Although it would be expected that E1 deleted viruses lack the machinery to express viral genes, several studies have demonstrated that cellular E1-like components can stimulate viral gene expression (Imperiale et al., Mol. Cell. Biol. 4:867-74, 1984; Onclercq et al., J. Virol. 62:4533-7,1988; Spergel et al., J. Virol. 66:1021-30, 1992). The expression of these viral genes results in the relatively rapid
25 elimination of transduced cells *in vivo* as a result of cytotoxic T cell responses (Yang et al., Immunity 1:433-42, 1994;. Yang et al., Gene Ther. 3:137-44, 1996; Yang et al., J. Virol. 69:2004-15, 1995).

Thus attention has been focused on eliminating the remaining vestiges of viral expression. Viral genes that have been deleted for this purpose include the gene for E4
35 proteins (Armentano et al., Hum. Gene Ther. 6:1343-53, 1995; Kochanek et al., Proc.

Natl. Acad. Sci. USA 93:5731-6, 1996; and Yeh et al., J. Virol. 70:559-565, 1996), DNA binding protein (Engelhardt et al., Proc. Natl. Acad. Sci. USA 21:6196-6200, 1994; and Gorziglia et al., J. Virol. 70:4173-8, 1996), DNA polymerase (Amalfitano et al., J. Virol. 72:926-33, 1998), and the preterminal protein (Schaack et al., Proc. Natl. Acad. Sci. USA 93:14686-91, 1996). The most aggressive approach has been the creation of helper virus-dependent vectors that lack all viral genes (Hardy et al., J. Virol. 71:1842-9, 1997; Kochanek et al., Proc. Natl. Acad. Sci. USA 93:5731-6, 1996; Lieber et al., J. Virol. 70:8944-60, 1996; Mitani et al., Proc. Natl. Acad. Sci. USA 92:3854-8, 1995; and Parks et al., Proc. Natl. Acad. Sci. USA 93:13565-13570, 1996). These vectors have high capacity, evoke reduced cellular immune responses and show prolonged expression *in vivo* (Morsy et al., Proc. Natl. Acad. Sci. USA 95:7866-71, 1998). However to deploy these viruses on the scale required for human clinical application presents major challenges because a cesium chloride (CsCl) gradient is needed to remove the helper virus.

15

Summary of the Invention

In one aspect, the invention features a replicatable viral DNA vector encoding a site-specific DNA-altering enzyme and a DNA target recognized by said enzyme, said enzyme selectively converting, in a cell expressing said enzyme, said DNA vector to a rearranged form.

20

In preferred embodiments, the rearranged form includes an autonomously replicating episome and a linear DNA product. In other preferred embodiments, the vector comprises adenoviral DNA.

In yet other preferred embodiments, the vector includes a genetically-engineered recombination site (such as a target of Cre or FLP). Preferably, such a recombination site includes a recognition sequence of a site-specific DNA altering enzyme.

25

In another preferred embodiment, the site-specific DNA altering enzyme is a recombinase (such as Cre or FLP) or an integrase. Preferably, such an enzyme is functional in a mammalian cell. Preferred embodiments of the vector also include an origin of replication that functions in a mammalian cell (such as an Epstein Barr Virus replicon). Moreover, the vector typically includes a gene of interest (such as a therapeutic gene that encodes a protein or polypeptide or an RNA product).

30

In another aspect, the invention features a method for assembling a recombinant adenoviral DNA. The method, in general, includes the steps of: (a) providing a first linearized DNA vector comprising a restriction site and a cos site and a second linearized DNA vector comprising the restriction site, an adenoviral nucleic acid molecule, and a cos site; and (b) ligating the first and second linearized DNA vectors, the ligation assembling a recombinant adenoviral DNA.

In preferred embodiments, the first linearized DNA vector comprises a selectable marker (such as a gene encoding a polypeptide that confers, on a host cell expressing such a polypeptide, resistance to an antibiotic). In other preferred embodiments, the first linearized DNA vector includes an adenoviral left-end inverted terminal repeat; a gene of interest; or both. In still other preferred embodiments, the second linearized DNA vector includes a selectable marker. Preferably, the second linearized DNA vector includes an adenoviral right-end inverted terminal repeat.

The method further includes packaging the assembled adenoviral DNA into a phage and infecting a host cell. Typically the first and second linearized DNAs include cosmid vector DNA. In addition, such adenoviral DNA is typically flanked by cleavage sites (such as intron endonuclease cleavage sites).

In another aspect, the invention features an adenovirus producer cell having a nucleic acid molecule that expresses a dominant negative site-specific DNA-altering enzyme. In preferred embodiments, the site-specific DNA altering enzyme is a dominant negative recombinase (for example, a Cre recombinase such as CreY324C or a Flp recombinase). Exemplary adenovirus producer cells include, without limitation, 293 human embryonic kidney cells, per.C6 cells, and N52 cells.

In yet another aspect, the invention features a vector comprising, in the 5' to 3' direction, a first genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; a gene of interest; a lineage-specific gene promoter; a second genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; and a nucleic acid molecule encoding a site-specific DNA altering enzyme.

In still another aspect, the invention features a vector including, in the 5' to 3' direction, a first genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; a gene of interest; a bi-directional promoter, comprising a second genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; and a nucleic acid molecule encoding a site-specific DNA altering enzyme.

In related aspects, the invention features a method of gene therapy including the administration to a patient in need of gene therapy a therapeutically effective amount of the vector of the invention, which is expressed in the patient. The invention further relates to a population of cells transfected with the vector of the invention.

5 Accordingly, the invention further relates to the use of a recombinant viral vector or use of a recombinant viral particle for gene therapy. Such vectors and viral particles may be introduced either *in vitro* into a host cell removed from the patient, or directly *in vivo*, into the body to be treated, according to standard methods known in the art.

 The invention also relates to a pharmaceutical composition that includes a
10 therapeutically effective amount of a recombinant viral vector or viral particle prepared according to the methods disclosed herein, in combination with a vehicle that is acceptable from a pharmaceutical standpoint. Such a pharmaceutical composition may be prepared according to the techniques commonly employed and administered by any known administration route, for example systemically (in particular, by intravenous,
15 intratracheal, intraperitoneal, intramuscular, subcutaneous, intratumoral, or intracranial routes) or by aerosolization or intrapulmonary administration.

 One skilled in the art will appreciate that suitable methods of administering a vector (particularly an adenoviral vector) of the present invention to an animal for purposes of gene therapy, chemotherapy, and vaccination are available, and, although
20 more than one route can be used for administration, one particular route may provide a more immediate and more effective reaction than another. Pharmaceutically acceptable excipients also are well known to those who are skilled in the art, and are readily available. The choice of excipient will be determined, in part, by the particular method used to administer the recombinant vector or particle. Accordingly, there are a wide
25 variety of suitable formulations for use in the context of the present invention.

 By "recombinant DNA vector" is meant a DNA sequence containing a desired sequence (such as a gene of interest) and an appropriate regulatory element(s) necessary for the expression of the operably linked sequence in a particular host organism (such as a mammal).

30 By "operably linked" is meant that a gene and a regulatory element(s) are connected to permit gene expression when the appropriate molecules (for example, transcriptional activator proteins) are bound to the regulatory sequence(s).

By "regulatory element" is meant a genetic element that controls some aspect of the expression of a nucleic acid sequence. For example, a promoter is a regulatory element that facilitates the initiation of transcription of an operably linked coding region. Other genetic regulatory elements include, without limitation, splicing signals, polyadenylation signals, and termination signals. For example, transcriptional regulatory elements in eukaryotes include promoter and enhancer elements. Promoters and enhancers include arrays of DNA sequences that interact directly or indirectly with cellular proteins involved in transcription. Promoter and enhancer elements have been isolated from a variety of eukaryotic sources including genes in mammalian cells and viruses.

By "transfection" is meant the introduction of foreign DNA into eukaryotic cells. Transfection is typically accomplished by a variety of means known in the art including, without limitation, calcium phosphate-DNA co-precipitation, DEAE-dextran-mediated transfection, electroporation, microinjection, liposome fusion, lipofection, protoplast fusion, and biolistics.

By "stably transfected" is meant the introduction of foreign DNA into the genome of the transfected cell. In general, transfer and expression of transgenes in mammalian cells are now routine practices to those skilled in the art, and have become major tools to carry out gene expression studies and to generate vectors useful in gene therapy.

By "gene of interest" is meant a gene inserted into a vector whose expression is desired in a host cell. Genes of interest include, without limitation, genes having therapeutic value, as well as reporter genes. A variety of such genes are useful in the invention, including genes of interest encoding a protein, which provides a therapeutic function. In addition, the gene of interest, if a therapeutic gene, can render its effect at the level of RNA, for instance, by encoding an antisense message or ribozyme, a protein which affects splicing or 3' processing (e.g., polyadenylation), or it can encode a protein which acts by affecting the level of expression of another gene within the cell (i.e., where gene expression is broadly considered to include all steps from initiation of transcription through production of a processed protein), for example, by mediating an altered rate of mRNA accumulation, an alteration of mRNA transport, and/or a change in post-transcriptional regulation.

By "reporter gene" is meant a gene sequence that encodes a reporter molecule (including an enzyme). A "reporter molecule" is detectable in any detection system, including, but not limited to, enzyme (e.g., ELISA, as well as enzyme-based histochemical assays), fluorescent, radioactive, and luminescent systems. Exemplary
5 reporter gene systems include the *E. coli* beta-galactosidase or glucuronidase genes, green fluorescent protein (GFP), blue fluorescent protein (BFP), the human placental alkaline phosphatase gene, the chloramphenicol acetyltransferase (CAT) gene; other reporter genes are known in the art and may be employed as desired.

By "transgene" is meant any piece of DNA, which is inserted by artifice into a
10 cell, and becomes part of the genome of the organism, which develops from that cell. Such a transgene may include a gene that is partly or entirely heterologous (i.e., foreign) to the transgenic organism, or may represent a gene homologous to an endogenous gene of the organism.

By "transgenic" is meant any cell that includes a DNA sequence, which is
15 inserted by artifice into a cell and becomes part of the genome of the organism, which develops from that cell.

By "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (for example, glycosylation or phosphorylation).

By "derived from" is meant isolated from or having the sequence of a naturally
20 occurring sequence (e.g., a cDNA, genomic DNA, synthetic, or combination thereof).

By "nucleic acid" is meant a polynucleotide (DNA or RNA).

By "gene" is meant any nucleic acid sequence coding for a protein or an RNA molecule.

By "gene product" is meant either an untranslated RNA molecule transcribed
25 from a given gene or coding sequence (such as, mRNA or antisense RNA) or the polypeptide chain translated from the mRNA molecule transcribed from the given gene or coding sequence. Nucleic acids according to the invention can be wholly or partially synthetically made, can comprise genomic or complementary DNA (cDNA) sequences, or can be provided in the form of either DNA or RNA.

30 The presently claimed invention affords a number of advantages. For example, applicants' gene therapy vehicles particularly those based on recombinant adenoviruses, minimize the propensity of the vectors to activate host immune surveillance, and thereby maximize the persistence for the DNA transduced. The invention therefore facilitates

the development of gene delivery vectors designed to enhance persistence of virally delivered genes and evade the cellular immune response by severing the connection between the sole adenoviral enhancer and the sequences encoding potentially antigenic viral proteins.

5 As described in more detail below, the mechanism by which this is accomplished differs significantly from any other previous approaches. For example, to reduce the immunogenicity of vectors it is widely acknowledged that some intervention, such as the removal of key genes, or the prevention of their expression in the cells targeted for therapy, is important; however, many related approaches are directed at the host and
10 have generally focused on the selective induction of tolerance to adenoviral antigens, or similar strategies directed at inducing a temporally restricted or antigen-specific compromise of the immune system.

 In addition, the poor persistence of transduced DNA appears to be due in part to immunological rejection of transduced cells and to the inability of the viral DNA to
15 replicate, a feature generally inherent in the design of adenoviral vectors, but one which is not associated with applicants' claimed gene therapy vehicles.

 Moreover, some contemporary adenoviral vectors are designed to propagate in specific host cells which provide essential replication factors *in trans*. These vectors are typically based on cell lines which express the master regulatory proteins of the E1
20 complex, which are required for induction of adenoviral DNA replication. In cells expressing E1 genes, the best studied of which is a human embryonic kidney cell line transformed by DNA from human adenovirus 5 (called HEK293, or simply 293), viruses lacking E1 genes propagate well. Such viruses do not propagate on cell lines which do not express E1, and do not generally propagate well in the target cells to which the
25 therapeutic gene is to be delivered. Cells transduced with E1-deleted adenovirus vectors also do not express high levels of viral genes in the absence of E1. However, the weak residual expression that remains in such vectors appears to be sufficient to induce cellular immune responses that contribute to the destruction of the transduced cells.

 In addition, the gene therapy vectors claimed herein are hybrid vectors capable of
30 self-rearrangement to form circular and linear DNA products. The linear DNA has a compromised ability to express adenoviral genes, and therefore has a lower immunological profile. And the circular DNA behaves like a mammalian plasmid, encoding the gene of interest and persisting by autonomous replication in the nucleus.

For example, the circularization of an adenoviral vector via the action of Cre recombinase beneficially places a gene of interest (for example, a therapeutic gene) on a self-replicating episome. Vector circularization occurs in a tissue-targeted manner, for example, as a result of the activation of a synthetic liver-specific promoter upstream of the recombinase Cre. Once circularized, the EBV replicon in the episome confers improved persistence on the therapeutic gene as detected by reporter gene expression and direct assay for the presence of vector DNA sequences.

Furthermore, the invention eliminates the requirement for a helper virus, thus avoiding two potential limitations of that system. First, the continuous expression of Cre recombinase may lead to toxicity in host cells, either as a direct consequence of the protein's activity or via its immunogenicity. Second, the Cre helper virus may itself produce antigenic viral proteins that contribute to the immunologic elimination of infected host cells. In contrast, the self-resolving adenovirus/EBV vector system disclosed herein advantageously provides no alternative source of viral proteins, and Cre expression is terminated upon rearrangement.

In addition, the invention described herein provides tools for analyzing the roles of the enhancer in viral gene regulation and virus growth.

The invention also provides a convenient general system for creating recombinant adenoviruses, which increase their attractiveness as gene transduction tools for basic research. The system, for example, employs two conventional plasmid vectors and a λ phage packaging step. The entire recombinant AdV genome is assembled into a single cosmid that is easily amplified in *E. coli*. The use of intron endonuclease recognition sequences flanking the ITRs enhances virus production while simplifying insertion of therapeutic gene sequences into the pLEP shuttle plasmid. The convenience of this vector system has facilitated the construction of over two hundred recombinant viruses to date.

Other embodiments and advantages of the invention will be apparent from the detailed description thereof, and from the claims.

Brief Description of the Drawings

FIGURE 1A is a schematic diagram of the structure of an adenoviral type A vector and its fate in a target cell. enh refers to the Ad2 enhancer; GFP refers to the marker gene green fluorescent protein; EBV refers to the Epstein Barr Virus replicon;

TetO₇ refers to a heptamer of Tet operator; TetR refers to the Tet repressor; VP16 refers to the viral protein 16 of Herpes simplex virus, SD refers to the splice donor site; and SA refers to the splice acceptor site.

FIGURE 1B is a schematic diagram of the structure of an adenoviral type B vector and its fate in a target cell. enh refers to the Ad2 enhancer; GFP refers to the marker gene green fluorescent protein; EBV refers to the Epstein Barr Virus replicon; SD refers to the splice donor site; and SA refers to the splice acceptor site.

FIGURE 2A shows a schematic diagram of the pLEP cosmid polylinker region and its position relative to the adenoviral left ITR. The adenovirus enhancer/packaging sequence (ψ) is boxed.

FIGURE 2B is a schematic diagram showing the generation of a single cosmid encoding the AdV genome by the direct ligation of two smaller plasmids. A gene expression unit, CMVGFP, was inserted into the pLEP cosmid at the polylinker region. pLEP and pREP cosmids were digested with an intron endonuclease (PI-PspI), ligated, and packaged *in vitro* to generate pAd2CMVGFP. This DNA was then digested with another intron endonuclease (I-CeuI) to expose the ITRs at both ends of the viral genome. Finally, cosmid digestion mixtures were transfected into 293 cells. Plaques generated by recombinant viruses are detected in 7-10 days.

FIGURE 3A shows the restriction analysis of cosmids carrying the full length AdV DNA showing uniform generation of the desired vector DNA. 2 μ g DNA samples from four pAd2-7CMVGFP colonies were digested with Bgl II, resolved on a 1% agarose gel and stained with ethidium bromide. The predicted sizes of the DNA fragments are: 13261, 7684, 5228, 5088, 2284, 1757, 1549, 1270, 351, and 275 base pairs (bp). The 5228 and 5088 fragments appear as a doublet, and the 351 and 275 bp fragments are too small to be seen on the gel.

FIGURE 3B shows the release of the recombinant Ad DNA from cosmids by I-CeuI digestion. 2 μ g of pAd2-7CMV DNA from two clones was digested with I-CeuI. Arrows indicated the position of the released recombinant AdV DNA and the vector fragments of approximately 35 kb and 5 kb, respectively.

FIGURE 4A shows the appearance of plaques in 293 cells transfected with 10 μ g of pIAdGFPB with no ITR exposed (undigested), one ITR exposed (BsaBI or I-CeuI), or both ITRs exposed (BsaBI plus I-CeuI). Values represent the mean plaque counts per

dish and the time required for plaque development in 293 cells from three separate experiments. "I" designates I-CeuI; and "B" designates Bsa BI.

FIGURE 4B shows the viral titers obtained from plaques that were allowed to grow over 10 days after transfection. Viruses were harvested and the titer of each virus stock was determined by a GFP based semi-quantitative titration procedure described
5 herein. Values represent the mean \pm SE of three independent determinations.

FIGURE 5 is a schematic diagram showing a linear AdV that resolves into a circular episome. The elements involved in the self-directed rearrangement of the vector are shown schematically in pLEP1BHCRGFP/EBV and in the corresponding AdV.
10 Starting from the left ITR, the elements are shown as following sequence: left ITR, 147 bp; first 34 bp loxP site; 185 bp enhancer/packaging signal; 64 bp splicing acceptor (SA) from EF1 α gene first intron; 720 bp GFP cDNA; 230 bp SV40 poly(A); 1.7 kb TK-EBNA-1/OriP; 970 bp HCR12 promoter; 1 kb EF1 α gene first intron containing splicing donor (SD) and acceptor (SA) sites with the second loxP site inserted at 64 bp upstream
15 of the 3' end; 1.2 kb Cre gene tagged with AU1 and a nuclear localization signal; ~120bp poly(A) signal and PI-PspI site. After infection of liver cells, the HCR12 promoter drives the expression of Cre which results in the cleavage of the two loxP sites. This results in the circularization of the fragment containing the EBV replicon. The excision severs the connection between the enhancer/packaging signals and the remainder of the
20 AdV genome. The Cre gene becomes promoterless and is left on the AdV genome fragment. After excision, the HCR12 promoter drives the expression of the GFP reporter gene. The EBV replicon maintains the excised circle as an episome in host cells.

FIGURE 6A is a schematic representation of the loxP sites and EBNA-1
25 locations in the AdV genome. The relevant Bgl II site is also shown.

FIGURE 6B shows the time course of rearrangement in HepG2 and Hela cells at an equal multiplicity of infection (moi) of 1,000 particles per cell. Cells were infected with Ad2HCRGFP/EBV viruses for 2 hours at 37 °C. Hirt DNA samples were extracted from the cells. ~5 μ g of Hirt DNA samples were digested with Bgl II, fractionated on a
30 1% agarose gel, and analyzed by Southern blot techniques using a ³²P-labeled EBNA-1 fragment as the hybridization probe.

FIGURE 6C shows the DNA blot results obtained from HeLa cells infected at a moi of 10,000; and HepG2 at 1,000. The upper bands (4915 bp) represent the circularized DNA fragments whereas the lower bands (3162 bp) represent the non-circularized AdV.

5 FIGURE 7A shows green fluorescent protein (GFP) expression in liver and non-liver cells infected with the Ad2HCRGFP/EBV viruses. Cells were cultured in 35 mm dishes and infected with the Ad2HCRGFP/EBV virus at desired moi. HepG2 cells were infected with 1,000 particles per cell, whereas HeLa cells were infected with a moi of 10,000. GFP expression was examined at the indicated time points after infection.

10 Fluorescent cells were photographed using an Olympus SC35mm camera mounted on an Olympus IX70 fluorescent microscope, at 200x magnification, using a filter with peak excitation and emission wavelengths of 450 nm and 510 nm, respectively.

FIGURE 7B shows the expression of GFP in HepG2, HeLa, A431, and HT29 cells. Cells were seeded in 35 mm dishes and infected with the Ad2HCRGFP/EBV virus at a moi of 10,000 particles per cell. GFP expression was examined at 72 hours after infection.

15

FIGURE 7C shows the expression of GFP in human primary hepatocytes. These cells were photographed under bright field (left) and fluorescent conditions (right).

FIGURE 8A shows the results of RT-PCR that was performed to detect the tripartite leader sequence (upper panel) for virus late gene expression; and PCR was performed in the DNA samples for detection of the AdV genomes. The specific target sequences are described in detail *infra*. PCR analyses of adenovirus late gene expression in cells infected with the first generation AdVs or the self-resolving Ad2HCRGFP/EBV was analyzed. HepG2 cells were cultured in 35 mm dishes and infected with increasing moi (0, 10, 100, 1000, 10,000, and 100,000) of adenoviral vectors. RNA and DNA were isolated in parallel from the cells at 72 hours after infection.

20

25

FIGURE 8B shows a summary of quantitative RT-PCR and PCR results. Each determinant was the average of three experiments.

30 FIGURE 9A is a schematic diagram depicting the deletion analysis of the OriP and EBNA-1 regions of the EBV replicon. Structures of the deletions in EBNA-1 and OriP are schematically represented. Elements considered important for episomal maintenance are indicated. FR refers to the family of repeats; DS designates the region

of dyad symmetry; LR1 refers to the so-called linker region 1; GA refers to gly-ala repeats; LR2 refers to linker region 2; and Dimerization designates the dimerization domain.

FIGURE 9B is a graph depicting fractions of GFP positive cells carrying the
5 EBV replicons represented in Figure 9A.

FIGURE 10A shows the positions and identities of Cre mutants tested for their dominant negative Cre activities.

FIGURE 10B is a schematic diagram of the substrate Cre plasmid (ad2239) used to test dominant negative functions of Cre mutants.

10 FIGURE 10C shows GFP expression in cells cotransfected with the substrate Cre plasmid (ad2239) and the indicated Cre mutants.

FIGURES 10D and 10E show Cre mutants tested for their ability to inhibit rearrangement. Only those showing the strongest inhibitory activities were retested in Fig. 10E. GFP intensity was normalized to that of cells in the absence of inhibition.

15 FIGURE 11 shows GFP expression in 293 TetON cells and #17 cells transfected with ad2239. The ability of #17 cells to inhibit Cre activity is demonstrated by the weak GFP signal in cells treated with 2 μ M doxycycline.

FIGURE 12 is a schematic diagram depicting the tetracycline mediated auto-regulatory circuit.

20 FIGURES 13A and 13B show the effects of different basal elements on synthetic TetO promoter activity. FIGURE 13A shows a schematic diagram of the components of various auto-regulatory synthetic TetO promoters. FIGURE 13B shows a comparison of the strength of auto-regulatory synthetic TetO promoters bearing different basal elements, in the presence and absence of tetracycline, using GFP as a marker in HepG2
25 cells.

FIGURE 14 shows the structure of a Cre substrate plasmid (ad2265). The promoter, $Ef1\alpha$, and the gene, BFP, are interrupted by two loxP sites, which can be joined by Cre-mediated recombination. PA stands for poly A; BFP for blue fluorescent protein.

30 FIGURES 15A and 15B show the estrogen regulation of Cre recombinase activity. 293 cells infected with type B virus, AD121.5, in which the Cre enzyme is fused with estrogen ligand binding domain at both the N- and C-termini were cultured in the presence or absence of 1 μ M estrogen. Cre-mediated rearrangement in the presence

of estrogen is shown in Figure 15A, whereas blot analysis of extrachromosomal DNA from the same cells is shown in Figure 15B. L represents the position corresponding to the unrearranged adenoviral DNA; and C represents the position corresponding to the circular form of DNA.

5 FIGURE 16 shows the rearrangement of adenoviral sequences *in vivo*. Extrachromosomal DNA from the livers of Rag-2 mice sacrificed 2.5 hrs post injection of type A adenovirus, AD102.7, was analyzed by DNA blot. L represents the size corresponding to linear adenoviral DNA; and C represents the size corresponding to rearranged circular DNA.

10 FIGURE 17 is a photomicrograph depicting high level GFP expression in Rag2 mouse hepatic tissues 48 hrs post type A adenovirus (AD102.7) injection.

 FIGURE 18A and 18B show schematic diagrams of the structures of adenoviral vectors and their fates in target cells. enh refers to Ad2 enhancer; GFP refers to green fluorescent protein; EBV refers to Epstein Barr Virus replicon; TetO₇ refers to heptamer of Tet operator; TetR refers to Tet repressor; VP16 refers to transcriptional activator domain from HSV protein 16; SD refers to splice donor site; and SA refers to splice acceptor site.

15 FIGURE 19A shows the structure of a FLP substrate plasmid, ad2879. The promoter, Ef1 α , and the gene, GFP, are interrupted by 2 FRT sites, which can be joined
20 by the FLP-mediated recombination. PA stands for poly A; BFP for blue fluorescent protein.

 FIGURE 19B shows the structure of a cre substrate plasmid, ad2204.

 FIGURE 20 shows the structures of several FLPe anti-sense plasmids.

 FIGURE 21 is a panel of photomicrographs showing inhibition of FLP enzyme
25 activity by anti-sense FLP. 293 cells were transfected with FLP substrate (Figure 12) and plasmids indicated in each photo. High GFP intensity indicate the higher expression of FLP and less inhibition by the anti-sense expressed.

 FIGURE 22 shows a schematic diagram of FRT/Cre and loxP/FLP adenovirus.

30

Detailed Description

Described herein are systems for the regulated self-rearrangement of DNA vectors, for example, gene therapy vectors. Such regulated self-rearrangement has the potential to prevent unwanted expression of vector genes not required for a therapeutic effect, and to allow the stable association of the therapeutic gene with the target cell.

The essential elements of the regulated DNA rearrangement system are a gene which encodes one or more proteins which induce DNA rearrangement, a method for regulating the activity of those proteins or their abundance, and a target DNA sequence on which those proteins act. Particularly desirable are methods for regulating the activity of the proteins or their abundance which can be easily carried out on an intact organism, such as administration or withdrawal of a drug, hormone, or environmental stimulus such as heat or irradiation, which induces the activity or abundance of the proteins which cause DNA rearrangement.

Especially desirable are regulated DNA rearrangement systems in which all of the components can be delivered in a single vector. An example of this is a virus which bears both the cis-acting sequences for DNA rearrangement as well as the protein or proteins which act on those sequences, and the regulatory apparatus which controls the activity or abundance of those proteins. However, it is not necessary that the different elements be encoded in a single nucleic acid.

The important elements of this strategy are: the compromise of vector gene function by regulated rearrangement of DNA topology, the generation of plasmid circles from vector DNA in a regulated manner, and the removal of enhancer or promoter elements from the vector DNA by regulated excision. It is also important that the circular DNA generated by site-specific recombination possesses a mechanism for stable association with the host genome in some form, here conferred by the EBV replicon. In other embodiments, the circular DNA might possess the ability to direct its integration into the host chromosomes by a site-specific integration. Site-specific integration into the host chromosomes may also be generated by the action of a regulated site-specific recombinase on a linear template without passing through a circular intermediate.

Also described herein is one particular self-rearranging vector that begins as a hybrid adenovirus vector which is capable of converting itself into two unlinked molecules, a circular and a linear DNA. After this event the linear DNA product is deleted for two important cis-acting sequences: the packaging signals, which are

required for insertion of the viral DNA into the viral capsid, and the enhancer, which increases the expression of other promoters encoded in the viral DNA. The remaining linear DNA is thereby compromised in its ability to express adenoviral genes, endowing the vector with a lower immunological profile. The circular DNA generated by the
5 excision event is a mammalian cell plasmid which has the capacity to persist by autonomous replication in the nucleus. This capacity is encoded in genetic elements derived from the Epstein Barr virus (EBV). A schematic diagram of such a vector is illustrated in Figure 5.

Epstein Barr virus is a human herpes virus which is the etiologic agent of
10 infectious mononucleosis and which has been implicated in the genesis of Burkitt's lymphoma, a B cell neoplasm, and is thought to be a predisposing factor for some forms of nasopharyngeal carcinoma. Approximately 85% of the adult Western population has a persistent population of B cells which contain a circular latent form of the viral-genome, maintained in cells by the action of Epstein Barr Nuclear Antigen 1 (EBNA1),
15 a DNA replication protein that acts on the viral latent phase origin of replication, OriP. EBNA1 in and of itself is not thought to promote neoplasia; current thinking places greater weight on the actions of the EBNA2 proteins and LMP, latent membrane protein, in the inception of EBV-associated neoplasm.

Mammalian cell plasmids have been created which bear the EBNA1 gene and
20 OriP. In nonrodent cells, these plasmids persist by replication with each transit of the cell cycle. Multiple transcription units can be borne by these plasmids, allowing regulated expression of diverse gene products.

Preferred adenoviral vectors, shown in Figures 1A and 1B, are linear forms of an EBV plasmid flanked by loxP sites, cis-acting sequences required for site-specific
25 recombination directed by the bacteriophage P1 cre protein. To prepare an adenovirus bearing both the cre protein and loxP sites, it is necessary to insure that the cre protein is not expressed while the vector is being propagated in 293 cells. To lower the immunological profile of the vector, it is also desirable that the cre protein not be expressed after the vector delivered its payload to the target cell and the cre protein
30 performed its function.

To accomplish these objectives, two general approaches have been developed for the production of adenoviral chromosomes that circularize following the regulated expression of site-specific recombinases. In each case, the vector is engineered to allow

for the production of viruses in 293 cells, and to provide transitory expression of recombinase that induces rearrangement in target tissues. The major difference between the two strategies lies in the way the deinduction of recombinase is achieved.

In the first approach, adenoviral vectors are engineered to turn an activating
5 transcription factor into a repressor upon chromosomal rearrangement. Vectors
employing this approach are referred to herein as type A vectors (Figure 1A). In the
second approach, the recombinase promoter is redirected following chromosomal
rearrangement. Vectors utilizing the second approach are referred to as type B vectors
(Figure 1B). In both cases a linear chromosome is converted to its circular episomal
10 form and a resulting deleted linear form. The circular DNA contains an Epstein Barr
virus (EBV) replicon, which allows synchronous replication of the episome with the host
mitotic cycle (Reisman et al., Mol. Cell Biol. 8: 1822-32, 1985; Yates et al., Nature 313:
812-15, 1985). The linear DNA is deleted for the enhancer and E1 genes.

One self-regulated gene switch, employing the type A vector strategy, was
15 designed based on the bacterial transposon Tn10 tetracycline repressor (tetR) gene. In
its natural context, the tetR protein binds to specific sequences (tet operator sequences)
upstream of a tetracycline resistance gene, preventing transcription of the gene unless
tetracycline is present. To adapt this protein for eukaryotic gene regulation, a gene
fusion is created between tetR and an active portion of a strong eukaryotic
20 transcriptional activator, the herpes simplex virus VP16 protein. The fusion protein
exerts its action on a synthetic promoter created by the insertion of multiple tet operator
sequences upstream of a basal promoter element. This configuration allows high-level
gene expression whenever the tetR-VP16 fusion protein binds to its cognate operator
sequences. Because the tetR protein normally does not bind to its operator in the
25 presence of tetracycline, the activity of this synthetic promoter is high in the absence of
tetracycline and low in its presence.

One example of a type A vector is shown in Figure 1A. This self-regulated gene
expression cassette, present in a hybrid adenovirus, consists of a bi-directional promoter
element in which central tetR binding sites are flanked by divergently oriented basal
30 promoter elements. In one direction the promoter directs the formation of a transcript
encoding the cre protein; in the other direction, the promoter directs the formation of a
tetR-VP16 fusion protein. The latter differs from the conventional version in bearing a
loxP site between the tetR component and the VP16 component. When tetracycline is

present this gene switch is silent. As shown in Figure 1A, upon introduction into a target cell in the absence of tetracycline, the tetR-loxP-VP16 fusion protein is produced, stimulating further production of the fusion protein, and the cre protein. The cre protein then acts to promote site specific recombination between the loxP site in the tetR-loxP-VP16 coding sequences, and a distant loxP site. As a result of this recombination, the fusion protein coding sequence is disrupted so that the promoter no longer directs the formation of a tetR-loxP-VP16 fusion protein, but gives rise to an inert tetR-loxP-VP16 fusion protein for binding to the promoter upstream elements, thereby extinguishing promoter activity.

As shown in Figure 1A, the excised circular DNA element contains at least two transcription units. In addition, other transcription units or internal ribosome entry site elements may be used to allow the coexpression of gene products which are useful for extending the persistence of the delivered DNA, regulating expression of the gene of interest, or providing for ablation of the transduced cells once their presence is no longer desirable. In addition, the linear DNA remaining after excision of the circular gene expression plasmid lacks both viral packaging sequences and the cis-acting enhancer. Within this linear DNA, additional loxP sites may be placed to provide for the rearrangement of the remaining vector DNA in the target cell, disrupting the normal topology of the genes, and further thwarting expression.

Using the type B vector design strategy, described in greater detail below, a recombinant adenoviral gene delivery system that is capable of undergoing growth phase-dependent site-specific recombination has also been constructed.

The following examples are presented for the purpose of illustrating, not limiting, the invention.

TYPE B VECTORS – EXPERIMENTAL RESULTS

Several experimental examples for constructing type B vectors and for carrying out the general approaches of the invention are now described below.

Two-Cosmid System for Efficient Construction of Recombinant AdV

To simplify and facilitate the generation of recombinant AdV, a system was established to assemble the desired AdV genome in a single plasmid by ligation (shown in Figures 2A, 2B). The system consists of two component vectors, a left end plasmid,

pLEP, and a right end plasmid, pREP. The left end Ad sequences (nt 1-376) in pLEP include the viral inverted terminal repeat, the cis-acting packaging sequences, and the viral enhancer. Nucleotide (nt) positions described herein refer to the wild type Ad2 sequence in GenBank (J019017). The Ad sequences are followed by the gene
 5 expression unit intended for delivery and an intron endonuclease (PI-PspI) cleavage site. The right end plasmid contains a PI-PspI site followed by the Ad2 genome from the end of the E1 locus rightward (nt 3527-35937).

pLEP is a small tractable vector for cloning, whereas pREP is much larger and contains less frequently manipulated genes. Both pLEP and pREP contain a
 10 bacteriophage λ *cos* site, oriented to generate a single cosmid of appropriate length for *in vitro* packaging following ligation of the two plasmids at the PI-PspI cleavage site. pLEP is tetracycline resistant (Tet^r) and pREP is ampicillin (Amp^r) resistant, allowing the recombinants to be selectively isolated by co-selection for both markers. In the resulting assembled cosmid, the adenoviral sequences are closely flanked by cleavage
 15 sites for the intron endonuclease I-CeuI. Digestion with I-CeuI liberates the entire recombinant AdV genome from the parent cosmid (see Figure 2B).

Three classes of pREP have been constructed to allow the preparation of AdVs bearing E1 (pREP7; SEQ ID NO.: 2), E1 and E3 (pREP8; SEQ ID NO.: 3), or E1, E3, and E4 (pREP12; SEQ ID NO.: 4) deletions. pREP7 (SEQ ID NO.: 2) contains nt
 20 3527-35937 of the Ad2 genome, and pREP8 (SEQ ID NO.: 3) carries an additional deletion in the E3 region (Δ nt 27901-30841). pREP12 (SEQ ID NO.: 4) has deleted open reading frames (ORF) 1-4 of the E4 region (Δ nt 34121-35469, 1348 bp). AdV generated with these cosmids should be able to accommodate 5, 8, and 10 kb inserts, respectively.

25 These aforementioned vectors were constructed as follows. The EcoRI to BsaI fragment that spans the ampicillin resistance gene in pBR322 was deleted and replaced by a synthetic adapter, and the bacteriophage λ *cos* site was inserted between the unique StyI and BsmI sites. A PCR amplified Ad2 fragment containing the left end ITR (L.ITR), enhancer elements, and the encapsidation signal (nt 1-376) was created and
 30 inserted into the adapter (Figures 2A, 2B) to yield the tetracycline-resistant left-end plasmid pLEP. The right end of Ad2 from the AflII site to the right end (nt 3527-35937) was assembled into an ampicillin resistant cosmid vector, pACKrr3 (SEQ ID NO.: 1), by multiple steps of PCR amplification and fragment interchange. The resultant cosmid

was termed pREP7 (SEQ ID NO.: 2). To expand vector capacity, two deletions were incorporated into the pREP7 (SEQ ID NO.: 2) cosmid, an E3 gene deletion (nt 27901-30841, 2840 bp); cosmid pREP8 (SEQ ID NO.: 3) and a 1.3 kb deletion (nt 34121-35469) in the E4 region of the Ad2 region; pREP12 (SEQ ID NO.: 4)

5 An example of the construction of an AdV carrying a CMV-GFP expression unit is outlined in Figure 2. pLEPCMVGFP (Tet^r) was digested with PI-PspI and ligated to the pREP7 (SEQ ID NO.: 2; ΔE1, Amp^r) digested with the same enzyme. The ligation mixture was packaged with λ phage extracts (MaxPlax lambda packaging extracts, Epicentre Technologies) and a fraction of the packaged phage was used to infect a
10 recombination-deficient *E. coli* host, with selection for the assembled plasmid on Amp/Tet plates. Transductants containing pLEP fused to pREP were selected on agar containing 25 μg/ml ampicillin and 12.5 μg/ml tetracycline (Amp/Tet). Colonies were selected and DNA isolated (Qiagen). DNA was used either for restriction analysis or for tranfection of 293 cells as described herein.

15 Figure 3A shows typical results for the Bgl II digestion pattern of a pLEP3CMVGFP/pREP7 hybrid cosmid, pAd2-7CMVGFP DNA. Because of the size minimum (~40 kbp) for λ phage *in vitro* packaging and the double antibiotic selection, most of the colonies growing on Amp/Tet plates were the desired hybrid cosmids, and undesired rearrangements were rarely seen. In the present example, all four pAd2-
20 7CMVGFP clones exhibited the digestion pattern predicted from the inferred sequence. The entire recombinant AdV genome was then released from the cosmid by I-CeuI digestion (Figure 3B). I-CeuI digestion leaves ten nucleotides to the left of the left ITR and eight nucleotides to the right of the right ITR. Short flanking sequences have been reported to be eliminated during replication of recombinant viruses after transfecting the
25 DNA into 293 (human embryonic kidney) cells (Hanahan et al., Mol. Cell. Biol. 4:302-309, 1984).

 The digestion reaction can be transfected into 293 cells without purification as follows. 293 cells, obtained from Microbix Biosystems (Ontario, Canada), were cultured in 10 cm dishes in complete Dulbecco's Modified Eagle's Medium (DMEM)
30 supplemented with 10% FBS, 2mM glutamine and penicillin/streptomycin (Gibco BRL), and maintained at 37 °C and 5% CO₂ atmosphere incubator. Cells were grown to ~50% confluence on the day of transfection. Ten μg of cosmid DNA were digested with I-CeuI in a volume of 50 μl. The reaction mixture was transfected into 293 cells by

calcium phosphate precipitation (Graham and Prevec, Manipulation of adenovirus vectors, p. 109-128, *In* E. J. Murray (ed.), Methods in Molecular Biology, vol. 7, Humana, Clifton, NJ, 1991) without purification. After transfection, cells were cultured and examined daily for the appearance of cytopathic effects (CPE). Virus propagation, purification, plaque assay, and viral DNA isolation were performed using established protocols (Graham and Prevec, *supra*). At day six post-transfection, 5-30 viral plaques/10 cm dish /10 μ g DNA were usually apparent, which compared favorably with the 30-50 plaques/10 cm dish /10 μ g DNA found for 293 cells transfected with purified wild type Ad2 DNA.

10 To compare the efficiency of recombinant virus production, similar viruses were also generated by homologous recombination. 20 μ g of pREP7 (SEQ ID NO.: 2) was co-transfected into 293 cells with 10 μ g of a plasmid encoding the left end of the adenoviral genome and a green fluorescent reporter gene (pLITREF1 α GFP). pLITREF1 α GFP contained the Ad2 left end nt 1-376, an EF1 α promoter/GFP expression unit and Ad2 sequence (from 3525-8120) that overlaps with the same sequence in pREP7 (SEQ ID NO.: 2). This overlap fragment served as the region for homologous recombination. Each co-transfection was performed in duplicate. Initial plaques took longer to appear (14 days post transfection) and were less abundant (0-3 plaques per plate).

20 Data in the literature suggest that exposed ITR ends favor efficient virus production (Hanahan et al., *supra*). To assess the importance of this effect, an AdV cosmid, pIADEF1 α GFPB, in which the AdV ITRs were flanked with a different restriction site at each end was constructed. pIADEF1 α GFPB DNA was digested with BsaBI to expose the right ITR, I-CeuI to expose the left ITR, or the two enzymes were used together to expose both ends. Digested cosmid DNA samples were transfected into 293 cells and plaques were allowed to develop. Virus propagation, purification, plaque assay, and viral DNA isolation were performed using established protocols described in Graham and Prevec. (Manipulation of adenovirus vectors, *In* E. J. Murray (ed.), Methods in Molecular Biology, vol. 7. Humana, Clifton, NJ., pp. 109-128, 1991).

30 Ten days after transfection the viruses were harvested and viral titers were determined. The average titer for the viral stocks (Figures 4A, 4B) was 1.3×10^4 pfu/ml from transfection with undigested DNA; 2.4×10^5 pfu/ml from BsaBI linearized DNA (free right ITR); 1.1×10^5 pfu/ml from I-CeuI linearized DNA (free left ITR); and $2.7 \times$

10⁶ pfu/ml for the BsaBI/I-CeuI double digested DNA (both ITRs free). Thus liberation of each end resulted in an approximate increase in the efficiency of generating virus by a factor of ten (Figures 4A, 4B).

5 Construction of an AdV Capable of Self-Rearrangement

One approach to attenuating adenoviral gene expression and improving transgene persistence is the creation of viruses capable of undergoing internal, self-directed rearrangement upon delivery to the target tissue. In principle, this objective can be achieved through the regulated expression of site-specific recombinases in vectors that
 10 contain the cis-acting target of recombinase action. To allow such vectors to be created, the recombinase activity must be suppressed during propagation in the packaging cell line. As described in more detail below, the use of a lineage-specific promoter to control recombinase expression has been successfully employed to achieve this end.

An example of this is shown in Figure 5. The expression of Cre recombinase
 15 was controlled by a liver-specific promoter constructed as follows. The human hepatic control region 1 and 2 (HCR1 and 2) of the ApoE/C gene locus (Allan et al., J. Biol. Chem. 270:26278-81, 1995; and Dang et al., J. Biol. Chem. 270:22577-85, 1995) were amplified by PCR using 293 cell genomic DNA as the template. The following primers were used to amplify both HCR1 and HCR2 fragment: HCRtop-
 20 5' gcggaattcggcttggtgacttagagaacagag 3' (SEQ ID NO.:5); HCRbot – 5' gcgggatccttgaacccggaccctctcacacta 3' (SEQ ID NO.:6). The amplified PCR fragments (~0.39 kb) were cloned into pUC19. The HCR1 and HCR2 sequences were confirmed by dideoxy DNA sequencing. The two fragments were assembled in a head to tail orientation, fused with a synthetic basal TATA element and cloned in a parental pLEP
 25 vector containing a GFP reporter gene. The resultant plasmid was named pLEPHCR12GFP. The synthetic liver-specific, as demonstrated below, provided a means to control Cre recombinase expression during propagation of the vector in 293 cells, and allowed for testing the consequences of abstracting the enhancer from the linear vector DNA upon delivery of the DNA to the target cells.

30 In 293 cells, this promoter is silent, allowing the viral chromosome to be propagated with minimal rearrangement. Any rearranged viruses that are formed lack packaging signals and so disappear from the pool of propagating vectors. In liver cells the Cre recombinase is induced by the action of the tissue-specific promoter. The

resulting Cre-induced recombination excises a circular episome and redirects the transcriptional output of the liver-specific promoter so that it directs the synthesis of the transgene of interest. The remaining linear fragment consists of an adenoviral genome lacking the enhancer and packaging signals and a Cre expression unit devoid of
5 promoter sequences.

In the form discussed here, one loxP site is located at nucleotide 147 of the Ad2 genome, between the left ITR and the enhancer/packaging sequences, and the second loxP site is placed inside an intron a few bases upstream of the splice acceptor sequence. Hence the loxP site does not appear in the resulting mature transcript. The Cre coding
10 sequence that remains on the right end linear fragment after rearrangement lies downstream from a splice acceptor that lacks a splice donor or upstream promoter sequences. This effectively terminates the expression of Cre following excision.

Prior to recombination, the Cre recombinase gene is under the control of a synthetic promoter (referred to as HCR12), consisting of hepatic locus control elements
15 from the human ApoE/C locus fused to the first intron of the human EF1 α gene. After cyclization the HCR12 promoter lies upstream of the transgene (in this case GFP) and the distal segment of the intron (beyond the loxP site) contains the adenoviral enhancer. To facilitate manipulation of the plasmids in *E. coli*, the human IgG1 hinge-CH2 intron (118 bp) was inserted in the Cre coding sequence at nucleotide 237, suppressing Cre
20 expression in bacteria. The circularized episome contains the latent origin of replication (OriP) and trans-acting DNA replication protein (EBNA-1) of Epstein Barr virus, and hence is capable of autonomous replication in synchrony with the host mitotic cycle (Yates et al., Nature 313:812-815, 1985).

Using the two cosmid system described above, the pLEP plasmid containing the
25 self-resolving components, pLEP1BHCR12, was ligated with pREP8 (SEQ ID NO.: 3; Δ E1 Δ E3) to create pAdVHCRGFP/EBV. The latter was digested with I-CeuI and transfected into 293 cells. Appearance of plaques from AdVHCRGFP/EBV was retarded (by 8 days) compared to non-rearranging viruses, perhaps as a result of basal expression of the liver-specific promoter in 293 cells. However high titer viral stocks of
30 10^{12} nominal (absorbance-determined) particles/ml was achieved.

Rearrangement in Target and Nontarget Cells

To test excision efficiency, HepG2 (hepatocellular carcinoma) and Hela (cervical carcinoma) cells, obtained from ATCC, were infected with virus at a multiplicity of infection (moi) of 1,000 nominal particles/cell. This titer corresponds to approximately 10 plaque forming units per cell. For these experiments, HepG2 and Hela cells were seeded in 35mm dishes and cultured to approximately 80% confluence in DMEM/FBS as described herein. Cells were infected with the desired multiplicity of virus in a volume of 1 ml at 37 °C for 2 hours. At the end of the incubation, cells were washed with PBS twice and cultured in 2 ml of medium. Cells were collected in parallel at desired points for low molecular weight DNA and RNA extraction. Cells were examined for GFP expression by fluorescence microscopy (Olympus, IX70) or microtiter plate reader (PerSeptive Biosystem, CytoFluor II) before extraction of DNA for analysis of chromosomal rearrangement.

DNA analysis of chromosomal rearrangement was performed as follows. 5 µg of Hirt DNA was digested with Bgl II and analyzed by DNA blot techniques using a labeled EBNA-1 gene fragment as probe (Figure 6). The Bgl II fragment from the non-circularized AdV is 3162 bp, generated from the 5' end of the AdV to the first BglII site in the AdV. The circularized fragment created from the two loxP sites has a size of 4915 bp (Figure 6A). Densitometry revealed that at 72 hours post infection, 95% or more of the input genomes had undergone circularization in HepG2 cells. In contrast, low but detectable levels of circularized fragment was visualized in Hela cells infected at the same time and at the same multiplicity of infection used for the HepG2 cells (Figure 6B).

At the time of infection (t=0, Figure 6B), the amount of input viral DNA detected by DNA blot was higher for HepG2 cells than for Hela cells when similar virus multiplicities were applied (moi of 1,000). This may reflect differences in AdV adsorption or infection efficiency between the two cell types, possibly as a result of the lower levels of coxsackievirus-adenovirus receptor on the Hela cells surface. To achieve similar viral genome input into HepG2 and Hela cells, Hela cells were infected with ten-fold more virus (moi of ~10,000) than HepG2 cells (moi of 1,000). Episomal DNA samples were extracted and analyzed by blotting. The results (Figure 6C) indicated that when comparable amounts of viral genome are present in the nucleus, the cyclization rate in both cell types was similar. Because the level of subsequent GFP expression is

much higher in HepG2 cells than in HeLa cells (Figure 7A), it is likely that very small amounts of Cre recombinase suffice to promote rearrangement, and that recombinase expression is not limiting for rearrangement in either HepG2 or HeLa cells.

5 GFP expression cannot be detected until rearrangement has taken place, so the measurement of the fraction of GFP positive cells provided a simple alternate method for assessing the degree of productive rearrangement. Figure 7A shows that GFP expression developed quickly in transduced HepG2 cells, but that only a few GFP positive cells can be detected in HeLa cells infected with a ten fold higher moi, conditions that allow circularization to a comparable extent to that seen in HepG2 cells
10 (Figure 6C).

The HCR12 promoter specificity was also tested by infecting two additional non-hepatic cell lines, A431 (human epidermoid carcinoma) and HT29 (human colon adenocarcinoma), with the Ad2HCRGFP/EBV vector. Both cell lines were obtained from ATCC and cultured using DMEM/FBS as described herein. A few cells, with
15 weak GFP signal, were detected at 72 hours after infection in these cells (Figure 7B). In contrast, these non-hepatic cells could be infected efficiently with a first generation AdV, Ad2CMVGFP virus (data not shown), indicating that the low GFP signal was not due to the low infectivity of these cells by AdV.

To further assess the utility of the AdV genome rearrangement, primary human
20 hepatocytes were infected with the Ad2HCRGFP/EBV vector. For these experiments, primary human hepatocytes, generously provided by Dr. Albert Edge (Diacrin, Inc., Charlestown, MA) were isolated and cultured as described by Gunsalus et al. (Nat. Med. 3:48-53, 1997), infected with adenovirus, and GFP expression was analyzed. As shown in Figure 7C, GFP expression was readily detected 72 hours after infection.

25

Diminished Viral Gene Expression in Rearranged AdV

After excision, the adenovirus major enhancer/packaging signal segregates with the episomal DNA, yielding a linear fragment containing the remainder of the AdV genome without this important cis-element (Figure 5). To assess the impact of enhancer
30 deletion, PCR amplification and quantitative RT-PCR measurement of late viral gene expression was performed as follows.

Four µg of total RNA was reverse transcribed into cDNA using M-MLV RT by a standard protocol (Promega). 1 µl of the cDNA from each sample was used in subsequent PCR reactions. PCR primers were designed to amplify the tripartite leader sequence of the adenovirus late genes: TPL1 - 5' act ctc ttc cgc atc gct gt 3' (SEQ ID NO.: 7) and TPL2 - 5' ctt gcg act gtg act ggt tag 3' (SEQ ID NO.:8). For detection of the AdV genome in the Hirt DNA samples, 1 µg DNA was employed in the PCR amplification using the following primers which are specific for the adenovirus DNA in the fiber gene: Fiber1 - 5' ccg cac cca cta tct tca ta 3' (SEQ ID NO.: 9) and Fiber2- 5' ggt gtc caa agg ttc gga ga 3' (SEQ ID NO.: 10). PCR reactions were performed as 95 °C 30 seconds; 54 °C 30 seconds; 72 °C 30 seconds for 30 cycles. All amplified products were analyzed on a 2% agarose gel.

For quantitative PCR, a molecular beacon based universal amplification and detection system was used (Intergen). A common leading sequence (Z sequence, 5' act gaa cct gac cgt aca 3') was added to the TPL1 and Fiber1 primers. The TPL2 and Fiber2 primers, described above, were used in the quantitative PCR reactions. 1 µl of the cDNA and one µg of Hirt DNA from each sample were used in the assay. The PCR were carried out in a 96-well spectrofluorometric thermal cycler (Applied Biosystems Prism 7700). The number of template molecules in the PCR reaction was calculated from the standard curve using linearized plasmid as templates.

As most late adenoviral genes transcripts share a common ~200 bp tripartite leader sequence (TPL) (Akusjarvi and Persson, Nature 292:420-6, 1981), the TPL sequence was chosen as a marker of viral gene expression. HepG2 cells were infected with the first generation vectors Ad2CMVGFP and Ad2HCRGFP, or the self-resolving vector, Ad2HCRGFP/EBV, using increasing multiplicities of infection. Total cellular RNA and low molecular weight DNA were isolated in parallel as described by Hirt (J. Mol. Biol. 26:365-9, 1967) and total RNA was prepared using RNazol solution (Tel-Test. Inc.). RT-PCR was performed to quantitate the amount of RNA encoding the TPL in the cDNA samples. PCR amplification of a 201 bp fiber gene fragment from the AdV genome was used to detect the amount of viral genome in the DNA samples. A representative result of three experiments is shown in Figure 8A. TPL sequences were detected, 72 hours post-infection, with either 100 or 1000 viruses infected per cell, using both of the first generation adenoviruses (upper panel).

In contrast, no TPL signal was detected in the self-resolving Ad2HCRGFP/EBV infected cells, even at a moi of 100,000/cell. PCR amplification of the AdV fiber gene revealed comparable levels of AdV genomic DNA in cells infected at comparable moi's. (Figure 8A, lower panel). The cDNA samples in which the TPL signals were detected were further analyzed by real-time fluorescence PCR. The corresponding genomic DNA samples were also analyzed to determine the number of AdV genomes present in each sample. The results are summarized in Figure 8B. There were approximately 1×10^4 TPL per 1×10^6 AdV genomes detected in the Ad2HCRGFP infected cells, but no detectable TPL in the self-resolving Ad2HCRGFP/EBV infected cells. These results indicate that adenoviral gene expression was dramatically reduced by the separation of the viral enhancer sequences occasioned by the re-arrangement of the self-resolving vector.

TYPE A AND TYPE B VECTORS – EXPERIMENTAL RESULTS

Additional experimental examples now follow that further illustrate the general approaches of the invention relating to using and constructing type A and type B vectors. For generating such adenoviral vectors, DNA sequences important for gene expression in the target tissue were placed between two loxP sites. The first loxP site was inserted between the Ad2 left-end inverted terminal repeat (ITR) and the enhancer sequence, replacing a BspLU11I and BstZ17 fragment of Ad2. A target gene expression cassette, comprising a promoter, a gene of interest, polyadenylation signals, the EBV replicon, and site specific recombinase expression unit were inserted in place of the E1 locus.

In type A adenoviral vectors, the second loxP site is placed between TetR and VP16, preserving the coding frame of both (Figure 1A). A bidirectional promoter in which a central heptamer of tetracycline operator sites (TetO) (Gossen and Bujard, Proc. Natl. Acad. Sci. USA 89:5547-5551, 1992) was flanked by two divergently oriented basal elements, directs the expression of TetR loxP VP16 from a synthetic TATA element, whereas Cre recombinase is controlled by the same heptamer of operator upstream of the HIV LTR basal element.

In the case of type B viruses (Figure 1B), the second loxP site was inserted in the first intron of the Ef1 α gene, which contains the transcription stimulating sequences described herein. In addition, a splice acceptor sequence was added to the 5' end of the coding sequence of the gene of interest. To avoid rearrangement during plasmid

construction in bacteria, the Cre recombinase coding sequence was interrupted by the addition of the human IgG1 hinge-CH2 intron (between amino acids Q78 and A79), as described herein.

5 Designing a Compact EBV Replicon

Most plasmids employing the EBV latent origin of replication exceed 10 kb in length. To provide a means for increasing the capacity of the recombinant adenoviral type A or type B vectors to accommodate a therapeutic gene, a compact EBV replicon having episomal stability was designed. To this end, deletions were generated in both
10 the cis-acting origin of replication, OriP, and the sequences encoding the trans-acting replication protein, Epstein Barr virus nuclear antigen-1 (EBNA-1) (Figure 9A). Episomal persistence was assessed with a green fluorescent protein (GFP)-bearing test plasmid by determining the fraction of cells retaining green fluorescence as a function of time, assuming that the half-life of GFP, in daughter cells that have not received an
15 episome as a result of segregation failure, is approximately 1.4 days (Fukumura et al., Cell 94:715-725, 1998).

EBNA-1 contains a central repeated structure that consists entirely of Gly and Ala residues, termed the GA repeats (Figure 9A). Although deletion of this structure has been reported to have little consequence, a deletion mutant consisting of both a short
20 OriP and a short EBNA-1 (SoriP + SEBNA1) was generated and found not to support plasmid maintenance effectively (approximately 40% loss per cell division). A version of this mutant, reconstructed with 40 GA repeats, in which the short OriP was paired with a short EBNA-1 provided significantly better plasmid stability (20% loss per cell division vs. 10 % per cell division for the wild type) (Figure 9B). Since most target
25 tissues are relatively quiescent mitotically, this level of segregation fidelity provides reasonable stability in a compact replicon.

Producing Cell Lines that Express Cre- or FLP-Dominant Negative Mutants

As discussed herein, one obstacle to creating adenovirus carrying both
30 recombinase and target sites has been the difficulty of controlling recombinase activity during virus propagation. Since efficient recombinase activity is needed in target cells, recombinase activity is best tempered in the production cell line.

Vector-independent methods to suppress recombinase activity during the production phase are attractive because they allow vector design objectives to be pursued with fewer constraints. In principle, dominant negative recombinase mutants provide the desired antagonism of recombinase activity. Cell lines expressing such
5 recombinase dominant negative mutants were produced as follows.

Dominant negative Cre mutants were selected from known point mutants (Wierzbicki et al., J. Mol. Biol. 195:785-794, 1987) that are defective in recombination function but are likely to retain dimerization function Figures 10A-E. Several mutants were screened for their abilities to inhibit Cre activity of a type B vector construct
10 (ad2239 in Figure 10B) in a transient cotransfection assay. Under these conditions, Cre activity is detected by the expression of GFP that occurs upon rearrangement. Figures 10D and 10E show the point mutants that were assessed and their relative activities in the transient cotransfection assay. Dilution studies, in which increasing amounts of substrate/Cre plasmid were cotransfected with the mutant forms, were conducted and
15 based on its favorable profile, one mutant recombinase, designated CreY324C, was chosen for further development (Figure 10D).

Strong constitutive expression of CreY324C, under control of the Efl α promoter failed to yield stable cell lines. Stable clones were obtained when the Efl α promoter was replaced with a tetracycline regulated promoter (Gossen et al., Science 268:1766-
20 1769, 1995). Clones were then tested for the ability to inhibit Cre enzymatic activities, and one clone, designated cell line #17, was selected for additional experiments. When a plasmid bearing Cre and capable of undergoing Cre-directed rearrangement to create a GFP transcription unit (ad2239) was transfected to #17 cells or parental 293ON cells, GFP expression in the #17 cells in the presence of 2 μ M doxycycline was significantly
25 lower than those of controls (Figure 11), showing that Cre enzyme activity can be inhibited in #17 cells.

In addition to dominant negative Cre mutants, dominant negative FLP mutants may also be identified. FLP belongs to the same family of site-specific recombinases as Cre recombinase. A number of FLP mutations that show defects in either cleavage or
30 ligation of FRT sites have been identified. Mutant FLP defective in cleaving FRT site (for example, H309L, L315P, G328R, G28E, N329D, S336Y, S336F, A339D, Y343F, and H345L) are generated using standard methods. Mutants that inhibit the wild type enzyme are then identified for generating stable cell lines according to the methods

described above. These and the other cell lines (described herein) are then used for producing FRT/FLP containing virus.

As mentioned above, difficulties creating stable cell lines expressing Cre dominant negative mutants were occasionally encountered. This difficulty is not limited to Cre mutants, but also to the wild-type Cre enzyme. In contrast, 293 cell lines stably expressing a thermostable FLP, referred to as FLPe (Buckholz et al., Nat. Biotechnol. 16:657-662, 1998), were created, suggesting that FLPe might not be as cytostatic as Cre protein. To demonstrate this, 293 cells were transfected with plasmids expressing either Cre or FLPe, and puromycin resistant colonies were selected. To generate stable cell lines expressing Cre or FLP mutants, 293 TetON cells were transfected with linearized plasmid expressing Cre or FLP mutants and puromycin acetyltransferase and selected with 1 µg/ml of puromycin. Puromycin resistant colonies were characterized further for their ability to inhibit Cre recombinase using the cre (ad2239) or flp (ad2879) substrate plasmids. Table 1 shows that there are more puromycin resistant colonies selected from FLPe transfected cells than from Cre transfected cells. From this result, it is expected that stable cell lines expressing a reasonably high level of dominant negative FLP may be readily created.

TABLE 1

Puromycin Resistant Colonies Formed When Cre Expressing or FLPe Expressing Plasmid was Used to Transfect 293 Cells

Expression Plasmid	Number of colonies (2 µg/ml puromycin)
Control	236
Cre	92
FLPe	127

Cre or Cre dominant negative mutants were also found to inhibit FLP activity (Table 2). Accordingly, cell lines such as cell line #17, that stably express a Cre dominant negative mutant (for example, CreY324C), are useful for producing FLP/FRT carrying adenovirus.

TABLE 2

Cre Inhibition of FLP Activity *in trans*

Plasmids	Arbitrary GFP intensity
Ef1 α FLP + FLP substrate + vector control	4.9
Ef1 α FLP + FLP substrate + Ef1 α Cre	0.78
Ef1 α FLP + FLP substrate + Ef1 α Cre R173C	2.23

FLP enzyme activity was measured by the GFP intensity by cotransfecting with a FLP substrate plasmid, ad2879 (Figure 19A). GFP intensity was quantified using IP lab software.

10 Transcriptional Regulation of Cre or FLP Recombinases

It has been relatively difficult to achieve high-level promoter inducibility in a replicating adenovirus. The challenge is similar to that of achieving faithful control of transcription in a transient expression setting. One approach to increase the induction ratio in a transient setting is the use of auto-regulatory (feed-forward) circuits. One such system, based on tetracycline dependent activation, is shown in Figure 12. A central heptamer of tetracycline promoter operator sites (TetO sites) was placed between two divergently oriented basal TATA elements. The leftward TATA controls the expression of the TetR-VP16 fusion protein, in which a loxP (or FRT) site has been placed between the TetR DNA binding domain and the VP16 transcriptional activator. The rightward TATA box directs the synthesis of recombinase, either Cre or the yeast FLP enzyme. In the presence of tetracycline, the promoter has reduced activity in both directions. Upon removal of tetracycline, the synthesis of both TetR-VP16 and recombinase are induced (Figure 1A). The induced recombinase then disjoins the TetR DNA binding element from the transcriptional activation contributed by VP16. Any existing TetR-VP16 fusions thereafter promote transcription of TetR, which competes with TetR-VP16 for TetO, resulting in deinduction of recombinase transcription.

When a model target cell line, HepG2, was tested with this type of adenovirus, the efficiency of circularization was low relative to that seen in 293 cells (data not shown), indicating a cell dependence of the bidirectional TetO promoter. To correct this, the TATA element of the TetO synthetic promoter (derived from the CMV immediate early promoter) was replaced with that of the HIV LTR. Constructs bearing differing components of the HIV basal promoter were analyzed for strength and regulation in 293 and HepG2 cells (Figure 13A). Among the constructs tested, one

version bearing the HIV LTR TATA and Sp1 elements (D in Figure 13A) showed the least basal expression in 293 cells (data not shown) and the greatest induction in HepG2 cells (Figure 13B).

Using this promoter, a construct (ad3400) containing the autoregulatory structure D as shown in Figure 13A was engineered, and Cre activities in the presence and in the absence of tetracycline were assayed. Plasmid ad2265 (Figure 14) in which a blue fluorescent protein (BFP) expression unit is interrupted by two loxP sites and transcription termination sequences was used as a substrate for Cre. Cre-mediated recombination joins BFP to the promoter resulting in BFP expression. As shown in Table 3, no difference was found in the intensity of BFP expression, either in the presence or absence of tetracycline. One possible explanation for this is that very little Cre protein is required for activity. Consistent with this idea, standard immunohistochemical techniques failed to reveal the presence of Cre enzyme in cells that were fully induced (data not shown).

TABLE 3

Cre Recombinase Activity Regulation in Type A Constructs

Construct	Cre Form	+tet -tam	-tet -tam	+tet +tam	-tet +tam
ad3400	Cre	2.58	3.39	2.71	6.20
ad4394	Cre-LBD	0.081	0.22	0.056	1.02
ad4705	LBD-Cre-LBD	ND	ND	ND	ND

Cre enzyme activity was measured in the presence or absence of the ligands, tamoxifen, by cotransfecting with the substrate plasmid (ad2265). BFP intensity (mean intensity/area) was quantified by analyzing fluorescent images captured by a digital camera using IP lab software. ND, refers to fluorescent intensities that were too weak to measure. Tet, Tetracycline; tam, tamoxifen.

Deletion of the PolyA Consensus Sequence from Cre or FLP Transcription Units

To reduce the expression of FLP or Cre recombinase further, the consensus polyA addition signals from the Cre or FLP transcript unit were deleted from vector constructs, leaving polyadenylation dependent on distal downstream sequences, for example, in gene IX. The activity of Cre using type B proviral constructs with or without the polyA signal was measured. As shown Table 4, the construct without polyA signals (AD229.3) showed a significant reduction of GFP intensity compared to a

construct bearing the polyA signal (AD230.5). When FLPe constructs of similar structure were evaluated, similar results were found (data not shown). These data show that Cre and FLPe enzyme activity levels can be modulated by attenuating polyadenylation.

5

TABLE 4

Effect of Deleting polyA Addition Signal From the Cre Expression Unit on Cre Enzyme Activity Level

10

	polyA	Relative GFP Intensity
AD229.3	-	0.25
AD230.5	+	1

Post-Transcriptional Regulation of Cre Recombinase Activity

Post-transcriptional control mechanisms of Cre recombinase activity were also evaluated. Translational fusions between Cre and the ligand binding domain (LBD) of estrogen receptor have been reported to be regulated by estrogen (Feil et al., Proc. Natl. Acad. Sci., U.S.A. 93:10887-10890, 1996; Gossen et al., Proc. Natl. Acad. Sci., U.S.A. 89:5547-5551, 1994), or, in the case of mutant estrogen receptors (Metzger et al., Proc. Natl. Acad. Sci. U.S.A. 92:6991-6995, 1995), by the partial antagonist tamoxifen.

Use of a ligand-dependent recombinase (ad4394 in Table 3), in combination with the HIV LTR-based autoregulated Tet system, allowed for a small degree of regulation by tetracycline, but not by ligand, as assayed using the ad2265 rearrangement assay (Table 3). One interpretation of this finding is that fusion of the estrogen receptor LBD to Cre provides only modest control of recombinase activity, but attenuates enzyme potency to a level so that transcriptional regulation can be measured.

To increase control of recombinase activity, the LBD was fused both to the N-terminus and C-terminus of Cre (LBD-Cre-LBD) and inserted into the coding sequence of both type A and type B vectors. When the LBD-Cre-LBD construct of type A was transfected into 293 cells, it showed no significant Cre enzyme activity even in the presence of ligand (Table 3). This result confirmed that the Cre recombinase activity is attenuated by N-terminal or C-terminal extension.

30

When the LBD fusion Cre enzymes were assayed in the type B vector context, only LBD-Cre-LBD fusions (pk8-ad4626) showed ligand-dependent regulation of Cre enzyme activities (Table 5). It appears that attenuated Cre activity in LBD-Cre-LBD, in the absence of ligand, is low enough to fall below the upper limit of the Cre assay.

5

TABLE 5

Cre Enzyme Activities of Type B Provirus

Cre Form	Provirus	-tam	+tam
Cre	pk8-ad2239	ND	ND
Cre-LBD	pk8-ad4332	4.1	6
LBD-Cre-LBD	pk8-ad4626	0.05	4.2

10

Cre enzyme activity was measured in the presence or absence of the ligands, tamoxifen. GFP intensity was quantified using IP lab software.
tam, tamoxifen.

15

Consistent with this notion, only the construct carrying two LBDs, pk8- ad4626, was able to produce virus (AD121.5) by transfection and propagate in 293 cells, while pk8-ad4332, which carried one LBD, produced virus (AD100.9) initially (following transfection of the cognate DNA) but was unable to propagate in 293 cells (Table 6). In the case of wild type Cre, no virus was produced in 293 cells by transfection.

20

TABLE 6

Production and Propagation of Type B Adenovirus

25

Cre Form	Type B adenovirus	Viral Production in 293 cells	Viral Propagation in 293 cells	Viral Propagation in #17 cells
Cre	Pack8-2239	-	-	-
Cre-LBD	AD100.9 (Pack8-4332)	+	-	+
LBD-Cre-LBD	AD121.5 (Pack8-4626)	+	+	+

30

The AD100.9 virus was able to propagate in #17 cells expressing the dominant negative Cre Y324C, demonstrating that modulation of Cre activity is important for viral production. Thus, adenovirus carrying both two loxP sites and Cre in two different configurations were generated by controlling Cre activity.

5

Viral Rearrangement in Culture

Cre/loxP mediated rearrangement of the adenovirus in tissue culture cells has also been analyzed. As shown in Figure 15A, the AD121.5 virus showed a significant increase in GFP expression in the presence of the ligand, estrogen, suggesting a successful rearrangement of the virus by Cre recombinases. When non-chromosomal DNA (Hirt, *J. Mol. Biol.* 40:141-144, 1969) was made from the cells and analyzed by DNA blot analysis, the viral DNA from estrogen treated cells was identified mostly in circular form (C in Figure 15B), while the DNA from cells not treated with estrogen was found mainly in linear form (L in Figure 15B).

To evaluate the efficiency of the self-rearranging viruses *in vivo*, high titer stocks of AD102.7 (a type A virus carrying LBD-Cre, pk8-ad4394) in #17 cells was prepared and purified by CsCl gradient ultracentrifugation. The titer of AD102.7 ($4-6 \times 10^{12}$ /ml by OD) is comparable to or slightly exceeds that of control viruses ($2-4 \times 10^{12}$ /ml by OD) which carry neither Cre nor a loxP site. To determine the efficiency of viral rearrangement *in vivo* and whether such rearrangement is dependent on the presence of ligand, AD102.7 virus (4×10^{11} pfu/mouse as determined by optical density) were injected via tail vein into Rag-2 mice that were pretreated with vehicle alone or 110 μ g/day of tamoxifen for 7 days as follows.

Rag-2 mice were injected with either PBS (mock) or 4×10^{11} adenovirus particles (as determined by OD₂₆₀) of type A virus, AD102.7, via the tail vein. At various times after injection, animals were sacrificed and the liver tissues were removed and frozen rapidly on dry ice. To visualize GFP expression in animal tissues, mice were anaesthetized and perfused with 4% paraformaldehyde containing 0.2% glutaraldehyde intracardially (Kafri et al., *Natl. Genet.* 17:314-317, 1997), and the liver tissues were removed and fixed overnight at room temperature in the perfusion buffer containing 30% sucrose. The fixed tissues were sectioned serially and observed under confocal scanning laser microscopy. In experiments evaluating the responses of ligand-regulated recombinase, mice were injected either with vehicle (vegetable oil) alone or with 110 μ g/day of tamoxifen for 7 days prior to adenoviral injection.

Liver tissues from these animals were harvested at 2.5 hrs post injection (the earliest time point taken after injection) and Hirt DNA from approximately 250 mg of frozen hepatic tissue was prepared and analyzed by blot analysis. As shown in Figure 16, the majority of adenoviral DNA was found in circular form in tissues from untreated mice, as well as tamoxifen-treated mice. It can be concluded from these data that the Cre enzyme activity present in the tissue, even in the absence of ligand was sufficient for efficient self rearrangement of virus. As expected, the hepatic tissues from the Rag2 mice injected with AD102.7 showed strong expression of GFP (Figure 17).

Demonstrating that AD102.7 virus, produced efficiently in 293 cells at high titres by the conventional means, can self rearrange efficiently *in vivo* provides the proof of the concept that potentially safer adenoviral gene therapy vectors can be produced.

Adenoviruses Carrying Both FRT and FLP Recombinase

Type A and type B proviral constructs carrying both FRT (FLP recombinase recognition site) and FLP recombinase were also generated. Structures of these viruses are analogous to those of loxP/Cre carrying viruses except that loxP sites are replaced by FRT sites and Cre coding sequence is replaced by FLP coding sequence (Figures 18A, 18B).

Virus Production at Reduced Temperature

Temperature dependence of the Ef1 α promoter using GFP expression as a marker was also examined. As shown in Table 7, Ef1 α promoter activity is strongly reduced at 32°C in comparison to 37°C or 39°C. The temperature sensitive nature of the Ef1 α promoter was used to propagate type B adenovirus carrying FLP at 32°C following initial production of the virus by DNA transfection (pk8-ad3302) at 37°C. HepG2 cells infected with these viruses (AD41.4) showed strong GFP expression, but with an approximately 12 hr delay compared to GFP expressing viruses, suggesting that FLP recombinase activity may be impaired at 37°C. To improve the activity of FLP recombinase, viral constructs were created using a thermostable FLP (referred to as “FLPe”) described by Buchholz *et al.* (Nat. Biotechnol. 16:657-662, 1998).

TABLE 7

Effects of Temperature on Ef1 α Promoter Strength as Shown by GFP Intensities

Tester plasmids	Arbitrary GFP Intensities			
		32°C	37°C	39°C
Ef1 α GFP	16 hrs	1475	7886	11409
	41 hrs	6472	36699	50787
	86 hrs	16256	53370	54424
Ef1 α Cre +ad2204	16 hrs	243	1141	2132
	41 hrs	1094	9119	9784
	86 hrs	695	3219	8144

5 GFP intensities were measured using a Fluorescent reader.

The activities of FLP and FLPe using a FLP substrate plasmid (ad2879, Figure 19A) in 293 cells were compared. As shown in Table 8, FLPe is significantly more
 10 active than FLP under these conditions.

TABLE 8

FLPe is Significantly More Active than FLP Recombinase

15

Plasmid		Mean GFP intensity
ad4821	Ef1 α FLPe	2.39
ad2949	Ef1 α FLP	0.01

Plasmid coding either FLPe or FLP was cotransfected with a FLP substrate plasmid (Figure 19A) into 293 cells. GFP intensity of each transfection was measured using IP lab program.

20

In addition, a tamoxifen-regulated FLPe was created by fusing the ligand-binding domain from a mutant form of estrogen receptor to the FLPe coding sequence at its C-terminus (FLPe-LBD). The FLPe-LBD was found to be regulated by the ligand, tamoxifen (Table 9). Although FLP activity was retained by C-terminal fusion (FLP-LBD), addition of a short oligopeptide tag to the N-terminus of FLP abolished its
 25 activity (data not shown).

TABLE 9

Tamoxifen Regulation of FLPe as Determined by GFP Intensities

Plasmid	FLPe	-tam	+ tam
ad4821 +ad2879	FLPe	++++	++++
ad5022 +ad2879	FLPe-LBD(tam)	+	+++

5

GFP intensity resulting from FLPe mediated recombination was measured using a fluorescent microscope. tam, 2 µg/ml tamoxifen.

10 Inhibition of FLPe Activities by Anti-sense FLPe

An anti-sense approach to inhibit FLP enzyme activity was also employed. This approach tested the notion that incorporation of an open reading frame into an antisense transcript would stabilize the transcript and potentiate antisense activity. Two approaches were utilized. In one approach, the BFP coding sequence was placed upstream of anti-FLPe. In the second approach, an anti-FLPe was placed upstream of an internal ribosome entry sequence (IRES) and the BFP coding sequence (Figure 20). The ability of these constructs to inhibit FLPe was assayed using a FLP substrate plasmid, ad2879 (Figure 19A) and the result is shown in Figure 21. These data show that anti-sense FLPe is more effective in inhibiting FLPe function when it is fused to BFP, which can presumably be replaced with any other stable protein.

20

OTHER SELF-REARRANGING ADENOVIRUSESMixed Infection With Adenoviruses Carrying loxP/FLP and FRT/Cre

One of the ways to produce adenoviruses that can be rearranged in target cells but not producer cells is to engineer two separate viruses, each carrying one recombinase and the target sequence for the other. To test this system, type B adenoviral constructs carrying Cre recombinase and FRT sites and FLPe recombinase and loxP sites were created (Figure 22). In target cells infected with both viruses, Cre catalyzes recombination between the two loxP sites in the FLP virus, and FLP carries out FRT mediated recombination in the Cre virus, resulting in two circular plasmids. The loxP virus contained BFP, whereas the FRT virus contained GFP. Measurement of the fluorescent intensities of GFP and BFP, after cotransfecting the two constructs, revealed that BFP expression (mediated by Cre enzyme) was greater than GFP expression

30

(mediated by FLP enzyme), suggesting that the Cre enzyme functions more efficiently than FLP.

Accordingly, these two recombinase activities in the target cells need to be balanced for complete circularization of both viral vectors. Exemplary methods for modulating Cre /FLP activity include the use of transcriptional regulation (such as by varying promoter strength and/or with or without poly A addition signal sequence) and translational and/or post-translational regulation (such as by changing FLP to FLPe and making LBD fusion proteins), and post viral production control (such as by changing the ratio of two viruses).

In one approach, Cre was replaced by Cre-LBD and FLP was replaced by FLPe. To improve identification of the rearrangement products, BFP was replaced with RFP as a marker for Cre recombination. As shown in Table 10, in the presence of estrogen, expression of GFP (FLPe mediated) and RFP (Cre mediated) were similar.

TABLE 10

RFP and GFP Expression of Cells Cotransfected With Type B Proviral Constructs Carrying Cre-LBD/FRT(GFP) or FLPe/loxP (RFP)

Plasmids	Genotypes	GFP intensity		RFP intensity	
pk8-ads120 +pk8-ads113	Cre-LBD/FRT (GFP) + FLPe/loxP (RFP)	Estrogen			
		-	+	-	+
		3222	3183	46	1954

To insure that both Cre and FLP carrying viruses with an optimal ratio infect each target cell, these viruses can be cross-linked prior to infection. For example, Cre carrying virus is labeled by biotin while FLP carrying virus is labeled by avidin. Mixing two types of modified viruses generates virus complexes of desired proportions as well. Biotinylation or avidinylation can be carried out using commercially available reagents such as EZ-Link TFP-PEO biotin (Pierce) and EZ-Link maleimide activated NeutrAvidin (Pierce). The extent of the biotin/virus and avidin/virus will be empirically determined to ensure the viability of the virus and to obtain an optimal ratio of two viruses in the complex. Optimal ratios will be those resulting in 1:1 Cre and FLP recombinase activities in target cells. The modifications will be done following manufacture's instructions.

This approach not only increases the effective capacity of adenoviral vector but also opens new avenue of applications involving multiple proteins, some of which cannot be coexpressed in production cell line as a result of combination toxicity.

- 5 All references mentioned herein are hereby incorporated by reference.
Other embodiments are within the claims.

What is claimed is:

10

Claims

- 5 1. A replicatable viral DNA vector encoding a site-specific DNA-altering enzyme and a DNA target recognized by said enzyme, said enzyme selectively converting, in a cell expressing said enzyme, said DNA vector to a rearranged form.
- 10 2. The vector of claim 1, wherein said rearranged form comprises an autonomously replicating episome.
3. The vector of claim 1, wherein said rearranged form comprises linear and circular DNAs.
- 15 4. The vector of claim 1, wherein said vector comprises adenoviral DNA.
5. The vector of claim 1, wherein said vector comprises a genetically-engineered recombination site.
- 20 6. The vector of claim 5, wherein said recombination site comprises a target of Cre or FLP.
7. The vector of claim 1, wherein said enzyme comprises a recombinase or an integrase.
- 25 8. The vector of claim 7, wherein said recombinase is Cre or FLP recombinase.
9. The vector of claim 1, wherein said enzyme is functional in a mammalian cell.
- 30 10. The vector of claim 5, wherein said recombination site comprises a recognition sequence of a site-specific DNA-altering enzyme.

11. The vector of claim 1, wherein said vector comprises an origin of replication functioning in a mammalian cell.

5 12. The vector of claim 11, wherein said origin of replication is an Epstein Barr Virus replicon.

13. The vector of claim 1, wherein said vector comprises a gene of interest.

10 14. A method for assembling a recombinant adenoviral DNA said method comprising the steps of: (a) providing a first linearized DNA vector comprising a restriction site and a cos site and a second linearized DNA vector comprising said restriction site, an adenoviral nucleic acid molecule, and a cos site; and (b) ligating said first and second linearized DNA vectors, said ligation assembling a recombinant adenoviral DNA.

15

15. The method of 14, wherein said first linearized DNA vector comprises a selectable marker.

20 16. The method of claim 14, wherein said first linearized DNA vector comprises an adenoviral left end-inverted terminal repeat.

17. The method of claim 14, wherein said first linearized DNA vector comprises a gene of interest.

25 18. The method of claim 14, wherein said second linearized DNA vector comprises a selectable marker.

19. The method of claim 14, wherein said second linearized DNA vector comprises an adenoviral right-end inverted terminal repeat.

30

20. The method of claim 14, said method further comprising packaging said assembled adenoviral DNA into a phage and infecting a host cell.

21. The method of claim 14, wherein said first and second linearized DNAs comprise a cosmid vector.

5 22. The method of claim 14, wherein said adenoviral DNA is flanked by cleavage sites.

23. The method of claim 22, wherein said cleavage sites comprise intron endonuclease cleavage sites.

10 24. An adenovirus producer cell comprising a nucleic acid molecule that expresses a dominant negative site-specific DNA-altering enzyme.

25. The producer cell of claim 24, wherein said site-specific DNA altering enzyme is a dominant negative recombinase.

15 26. The producer cell of claim 25, wherein said recombinase is a Cre or Flp recombinase.

20 27. The producer cell of claim 26, wherein said dominant negative recombinase is CreY324C.

28. The producer cell of claim 26, wherein said Flp recombinase is Flpe.

25 29. The producer cell of claim 24, wherein said cell is a 293 human embryonic kidney cell.

30. A vector comprising, in the 5' to 3' direction,
a first genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme;

30 a gene of interest;

a lineage-specific gene promoter;

a second genetically engineered *cis*- acting target recognized by a site-specific DNA altering enzyme; and

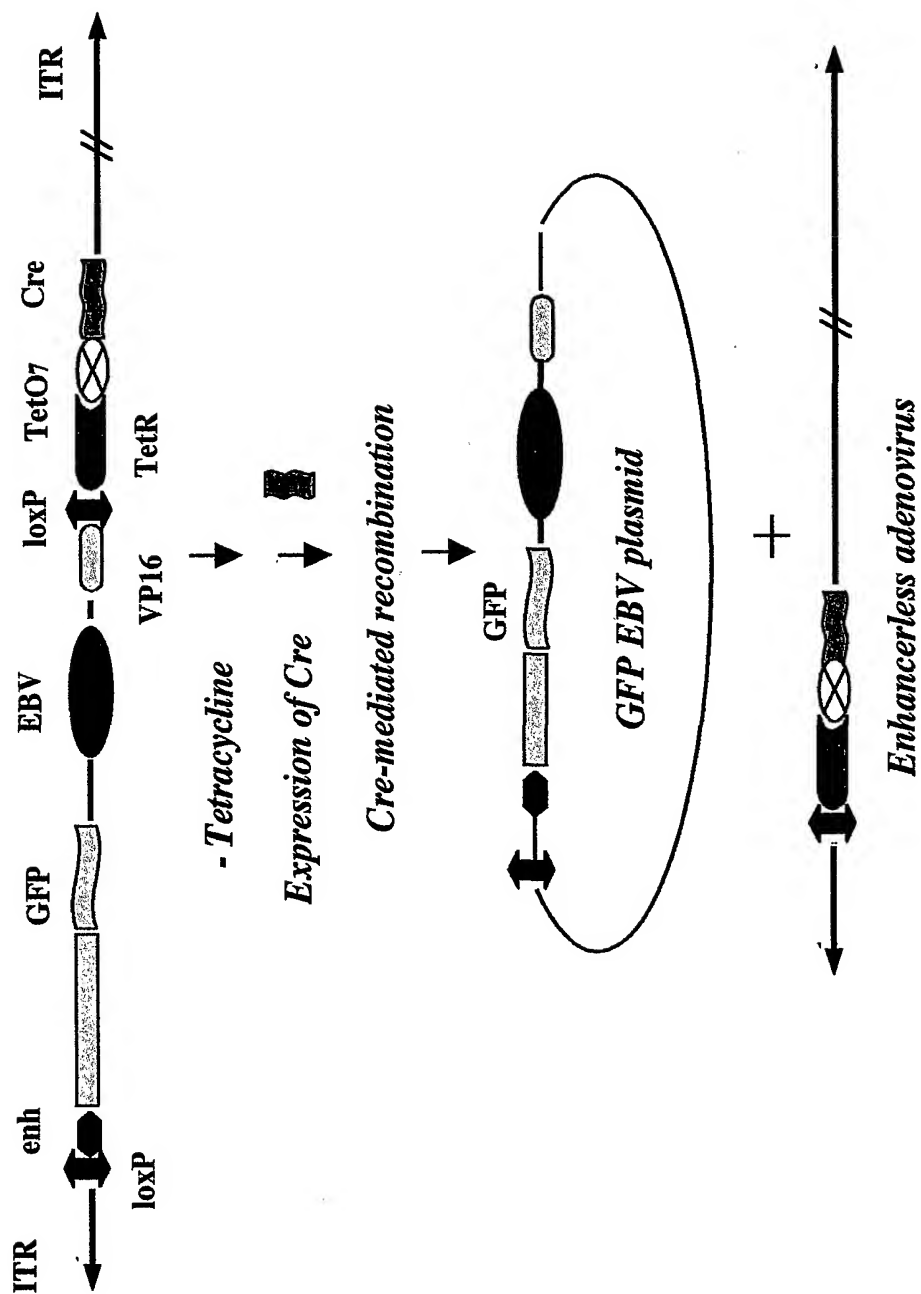
a nucleic acid molecule encoding a site-specific DNA altering enzyme.

35

31. A vector comprising, in the 5' to 3' direction,
a first genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme;
a gene of interest;
5 a bi-directional promoter, comprising a second genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; and
a nucleic acid molecule encoding a site-specific DNA altering enzyme.
32. A method of gene therapy comprising the administration to a patient in
10 need of gene therapy a therapeutically effective amount of the vector of any one of claims 1, 30, or 31 which is expressed in said patient
33. A population of cells transfected with the vector of any one of claims 1,
30, or 31.
15
34. A method of gene therapy comprising the administration to a patient in
need of gene therapy a therapeutically effective amount of the population of cells
of claim 33.
20

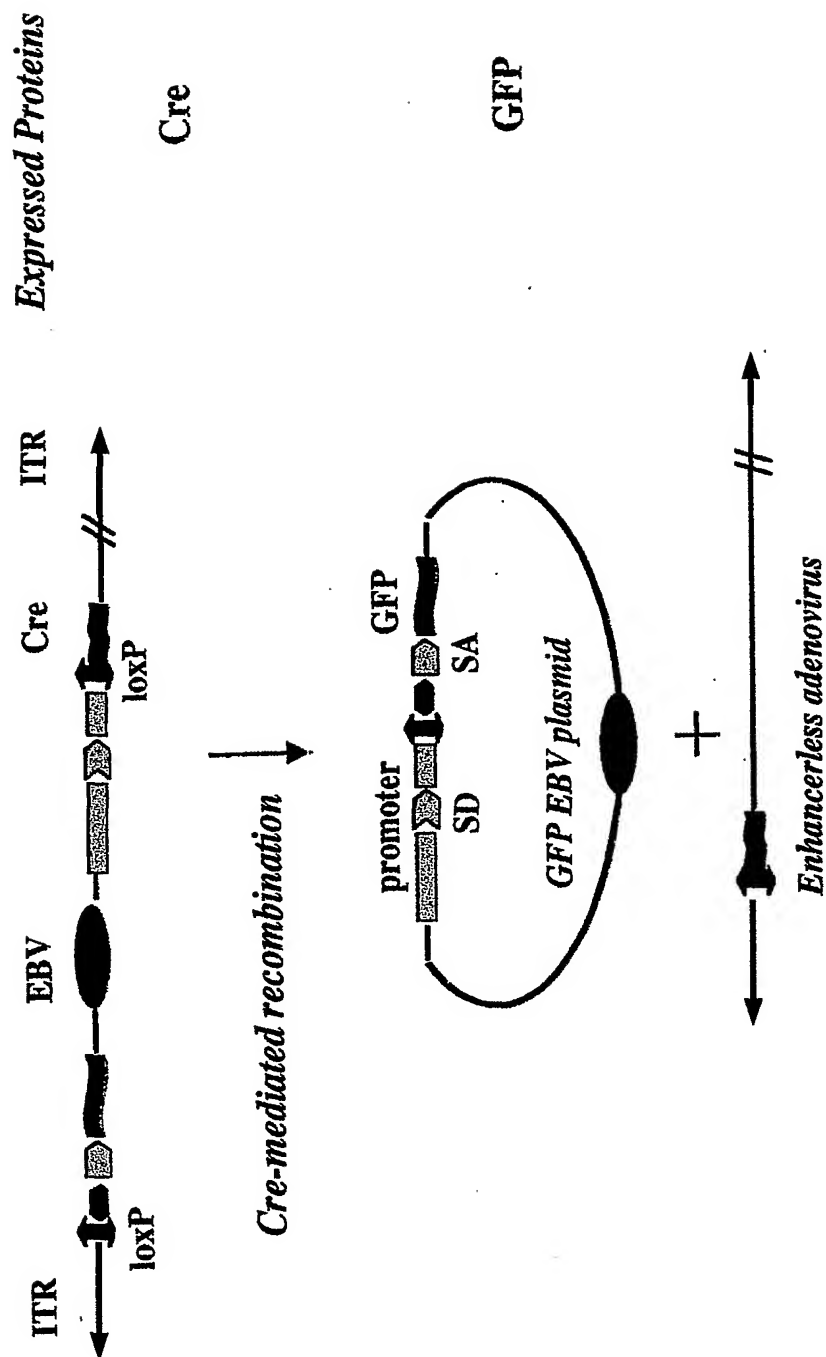
1/40

FIG. 1A



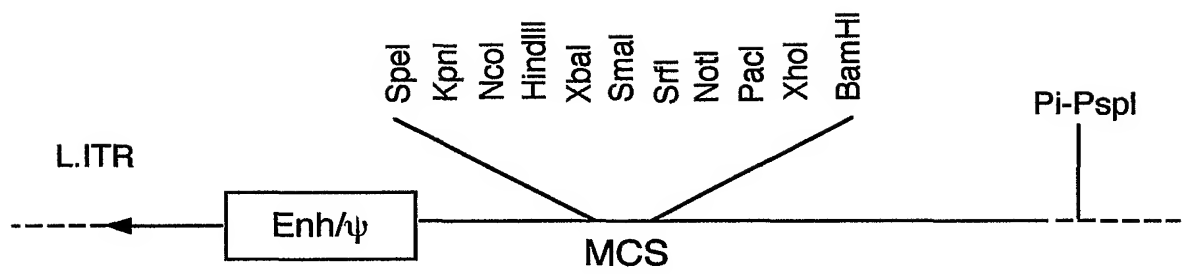
2/40

FIG. 1B



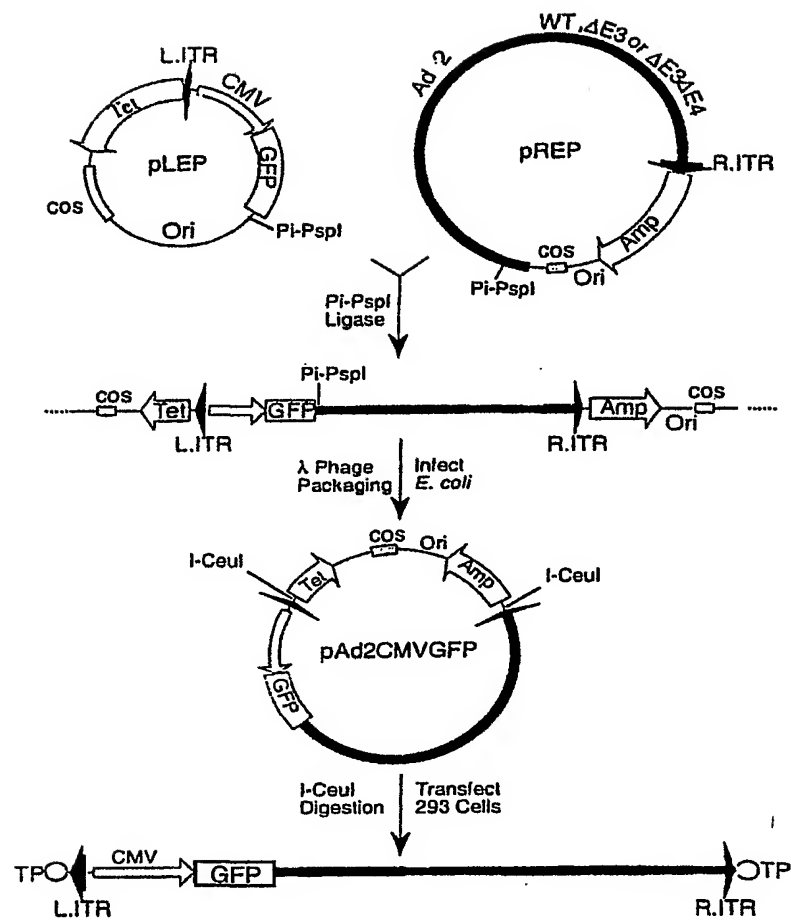
3/40

FIG. 2A



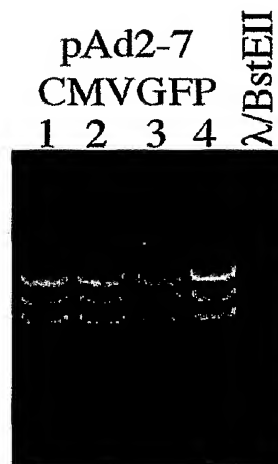
4/40

FIG. 2B



5/40

FIG. 3A



6/40

FIG. 3B

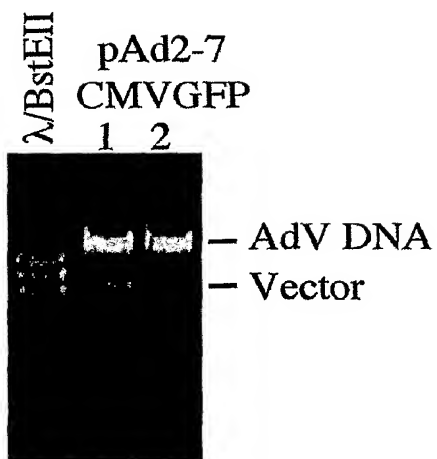
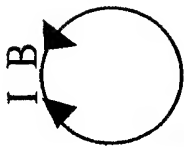



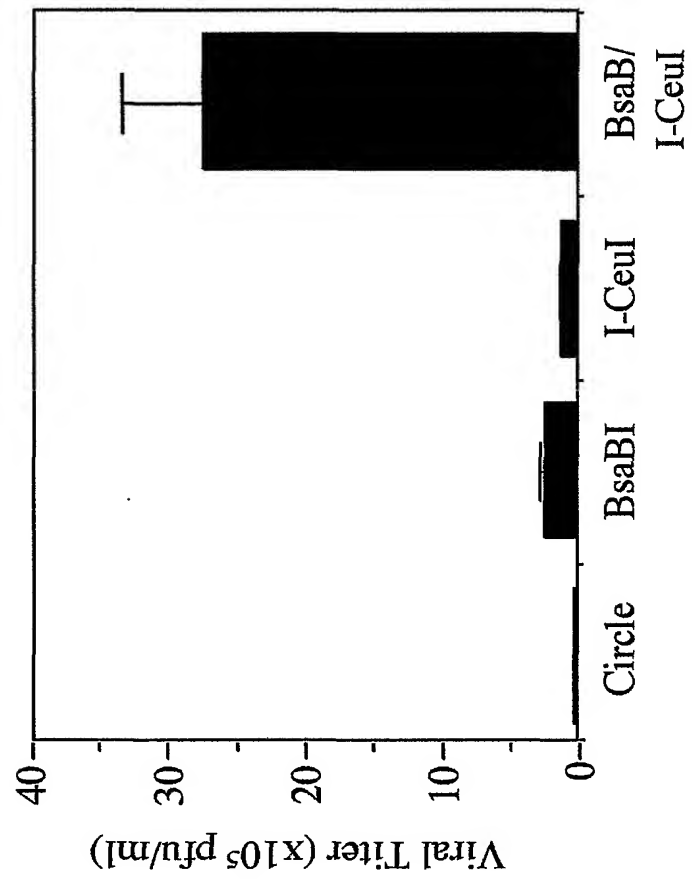


FIG. 4A

Ad DNA forms		Digestion	CPEs	Time (days)
1. Circular Ad DNA		pIAd2B	0.17	--
2. Linearized cosmid DNA		pIAd2B / I-CeuI	8.5	10
		pIAd2B / BsaBI	8.5	10
3. Liberated AdV DNA from cosmid		pIAd2B / I-CeuI + BsaBI	76	6

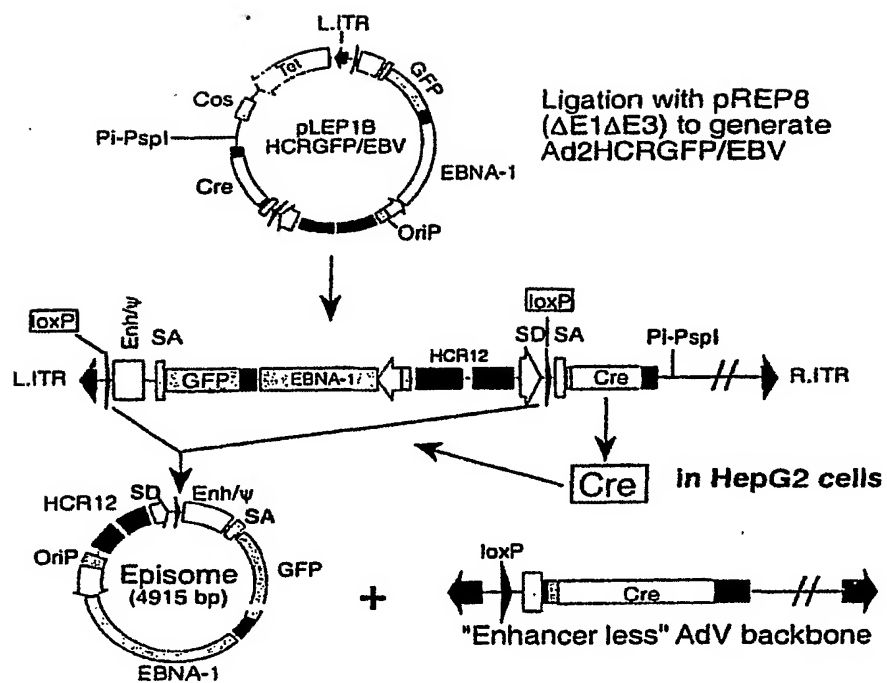
8/40

FIG. 4B



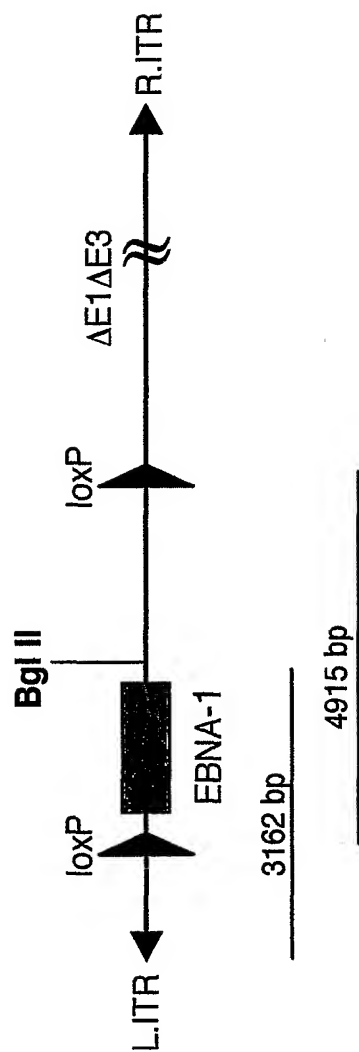
9/40

FIG. 5



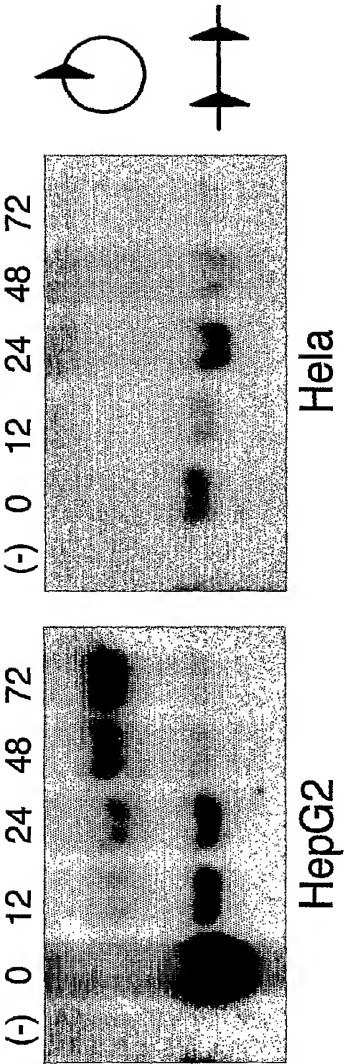
10/40

FIG. 6A



11/40

FIG. 6B



12/40

FIG. 6C

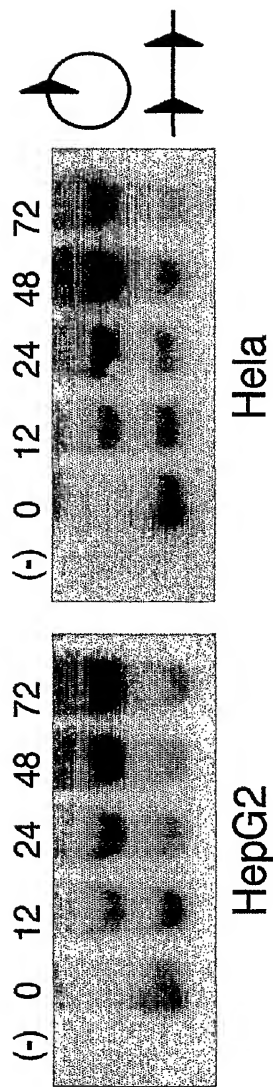
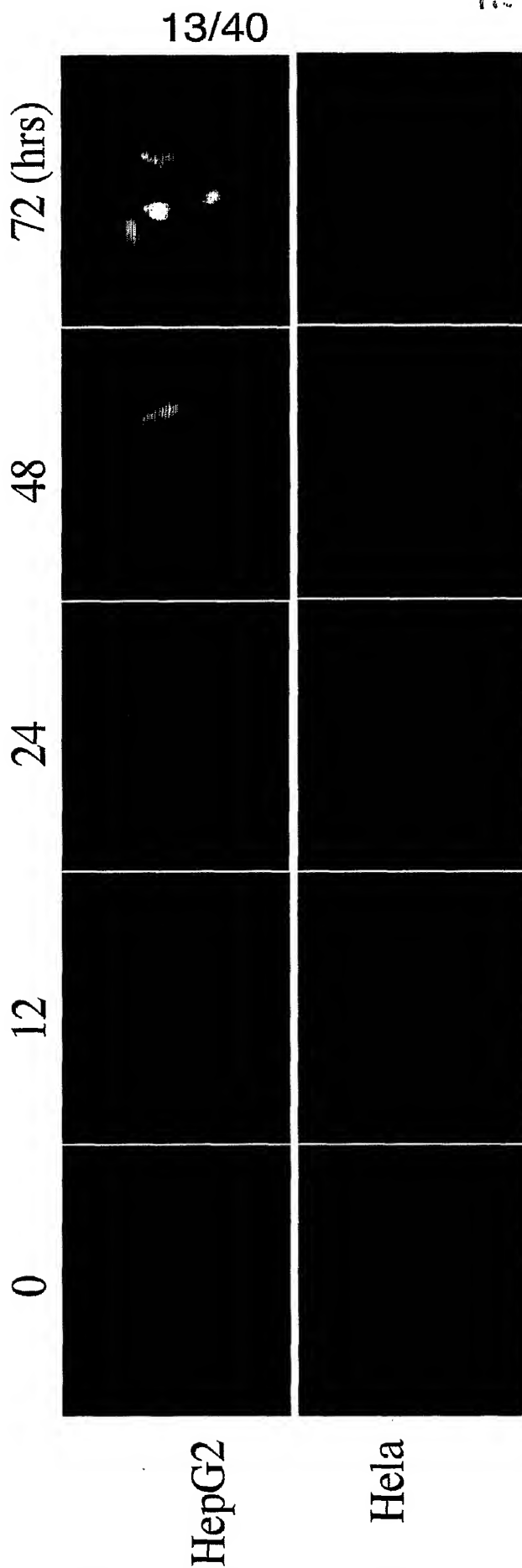
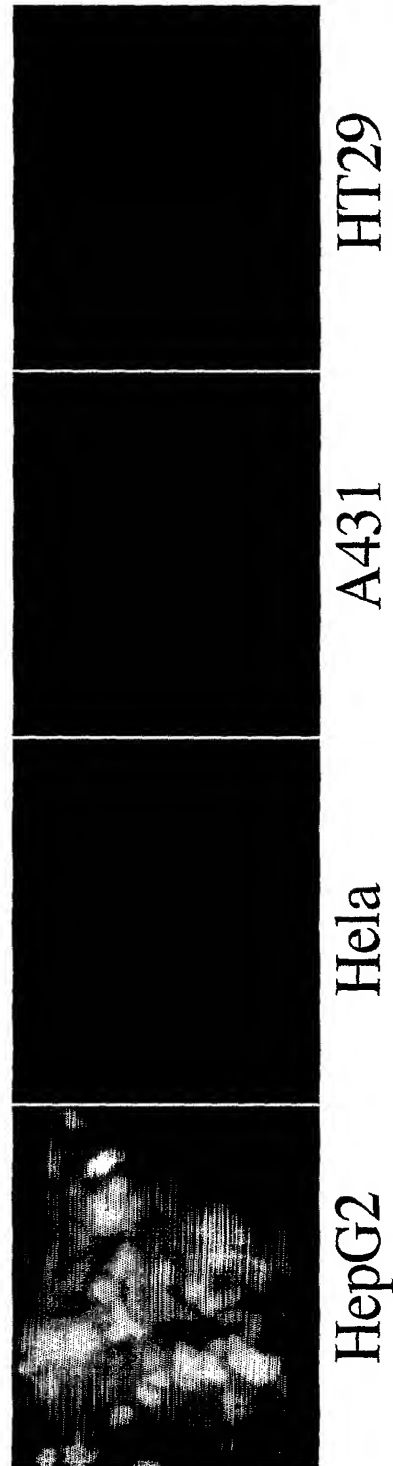


FIG. 7A



14/40

FIG. 7B



15/40

FIG. 7C



Human primary hepatocytes

FIG. 8A

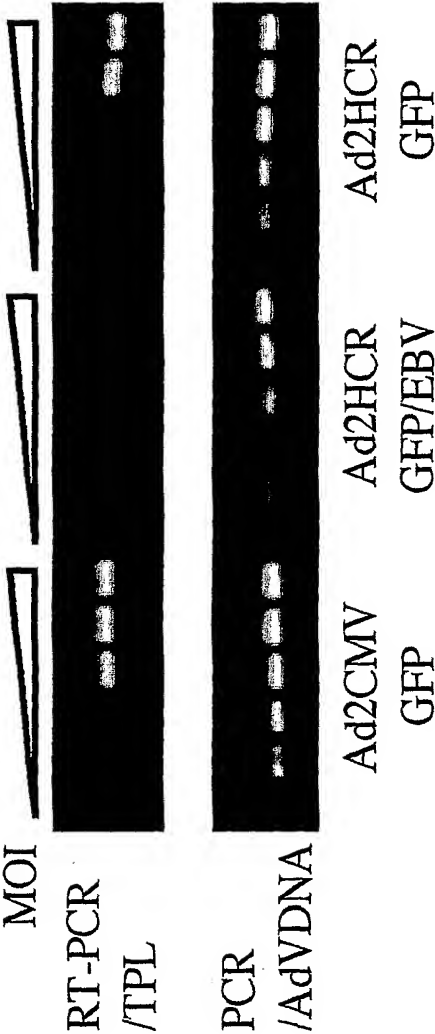
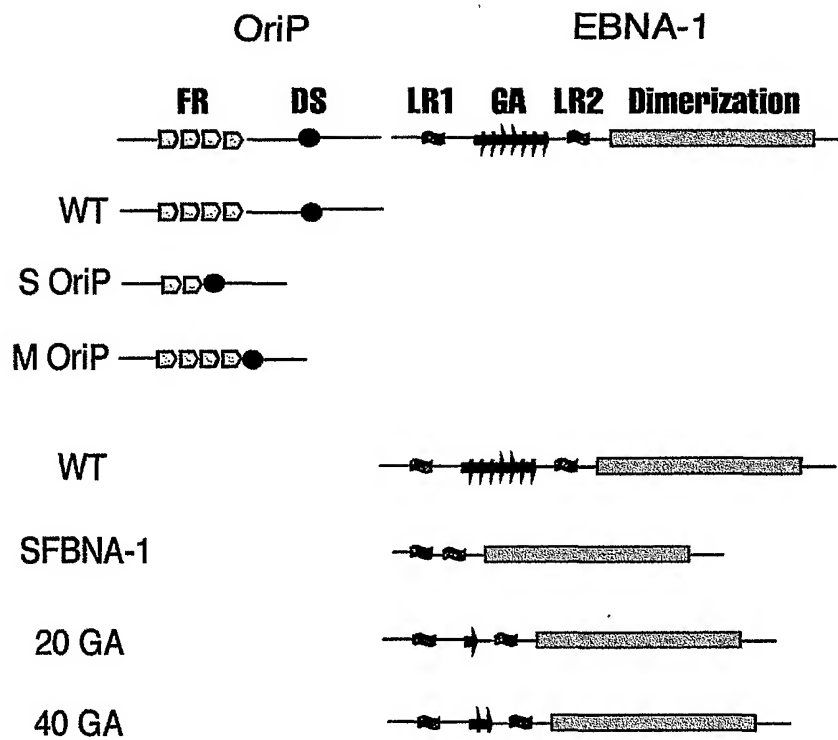


FIG. 8B

	# of TPL mRNA / 10 ⁶ AdV DNA
Ad2HCRGFP	1.15 ± 0.28 x 10 ⁴
Ad2HCRGFP/EBV	Non-detectable

18/40

FIG. 9A



19/40

FIG. 9B

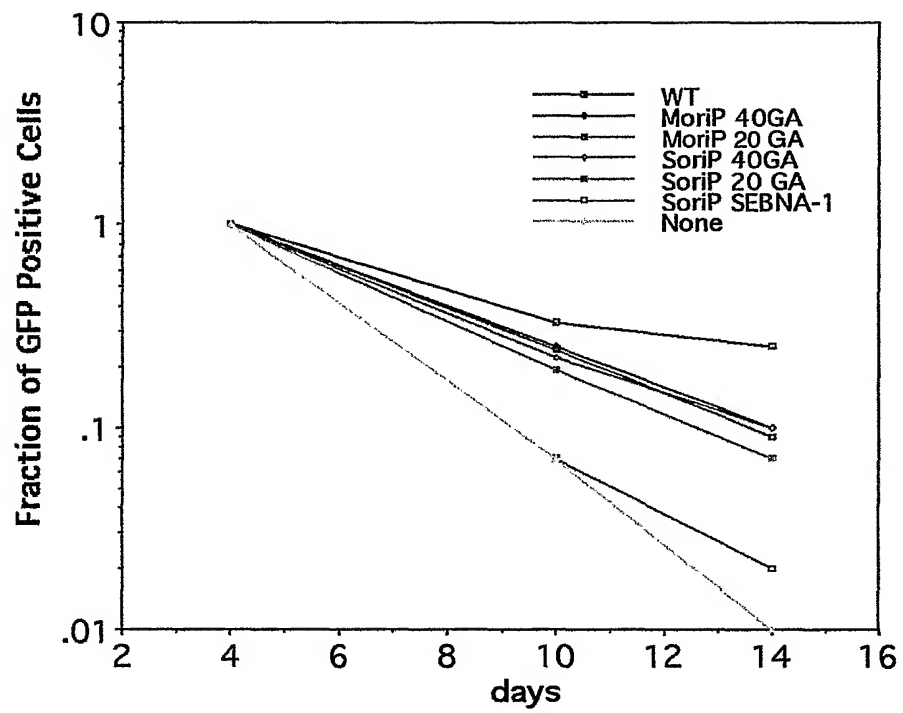
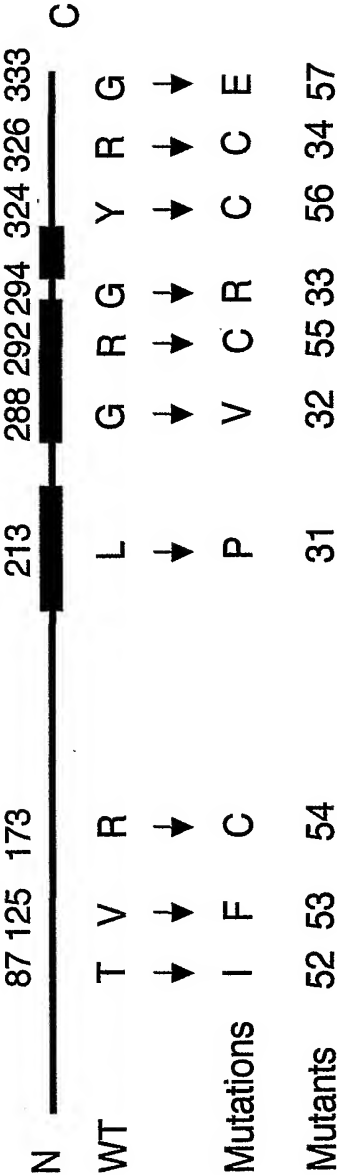
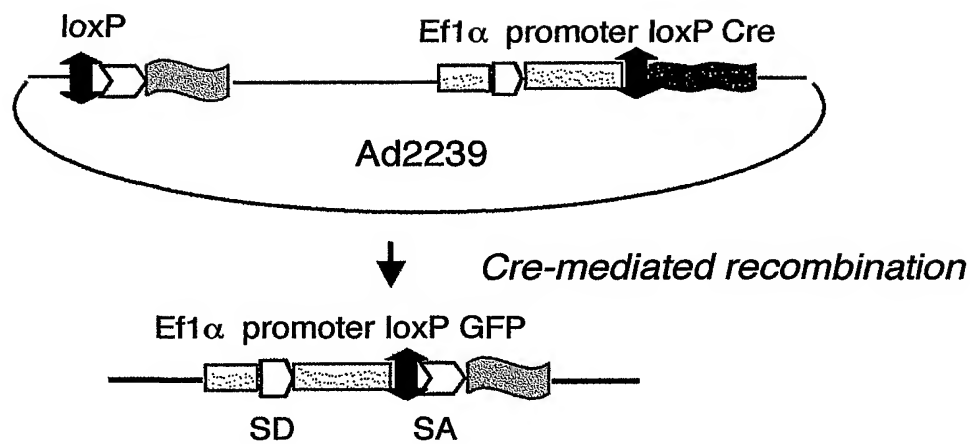


FIG. 10A



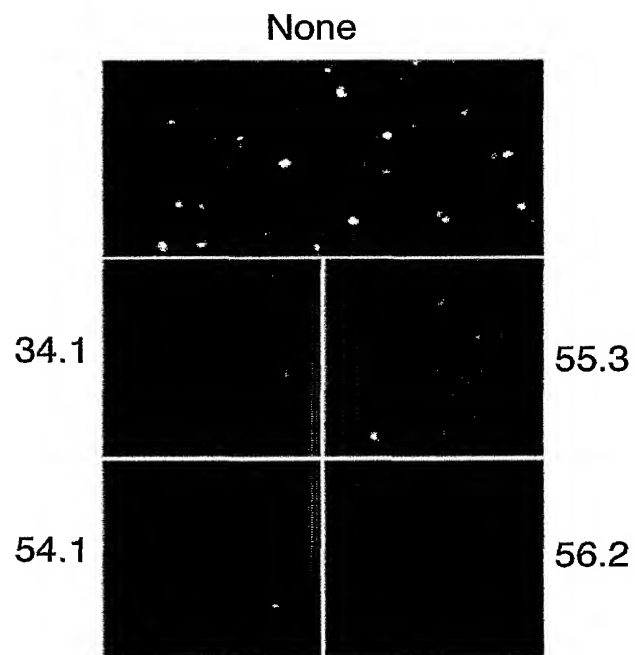
21/40

FIG. 10B



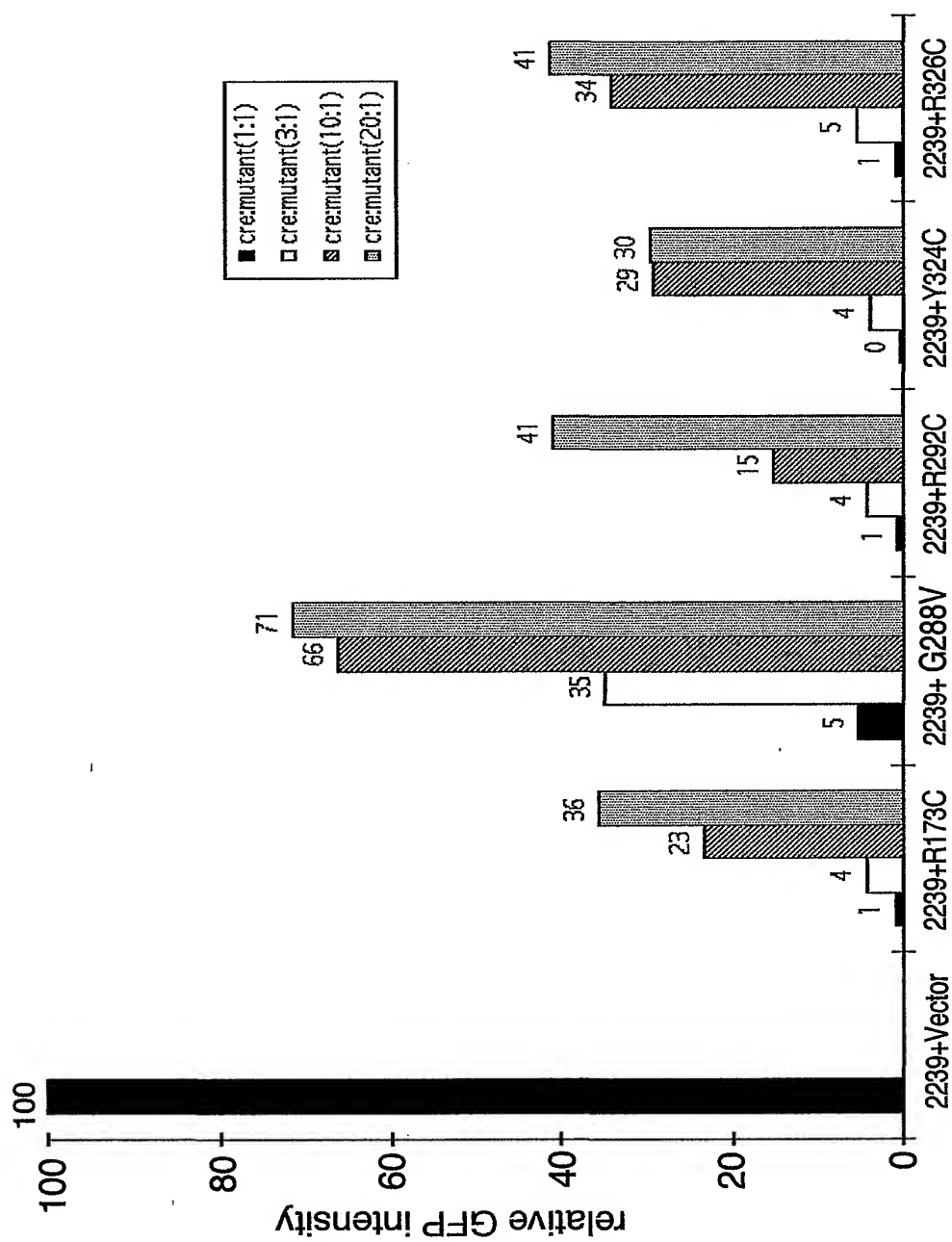
22/40

FIG. 10C



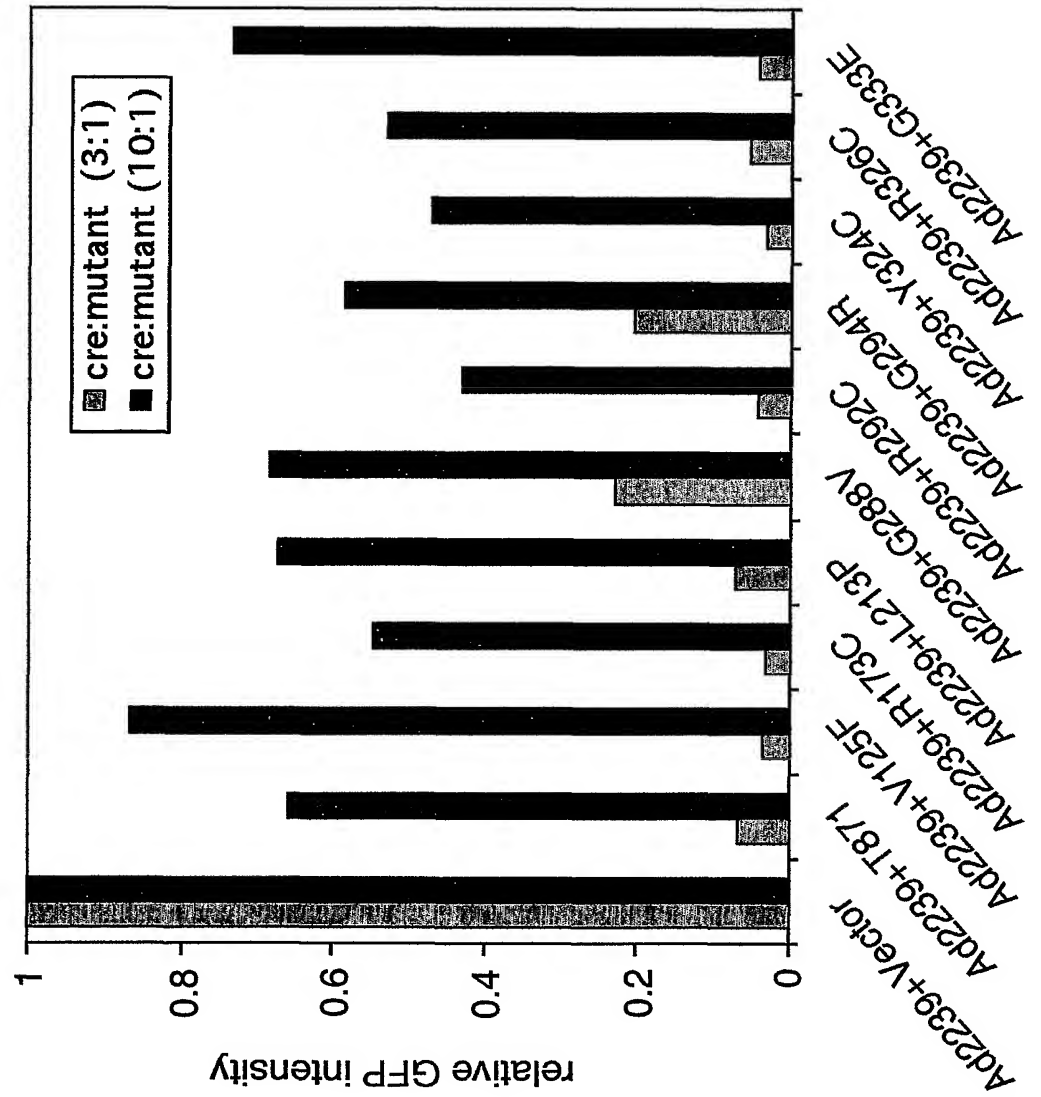
23/40

FIG. 10D



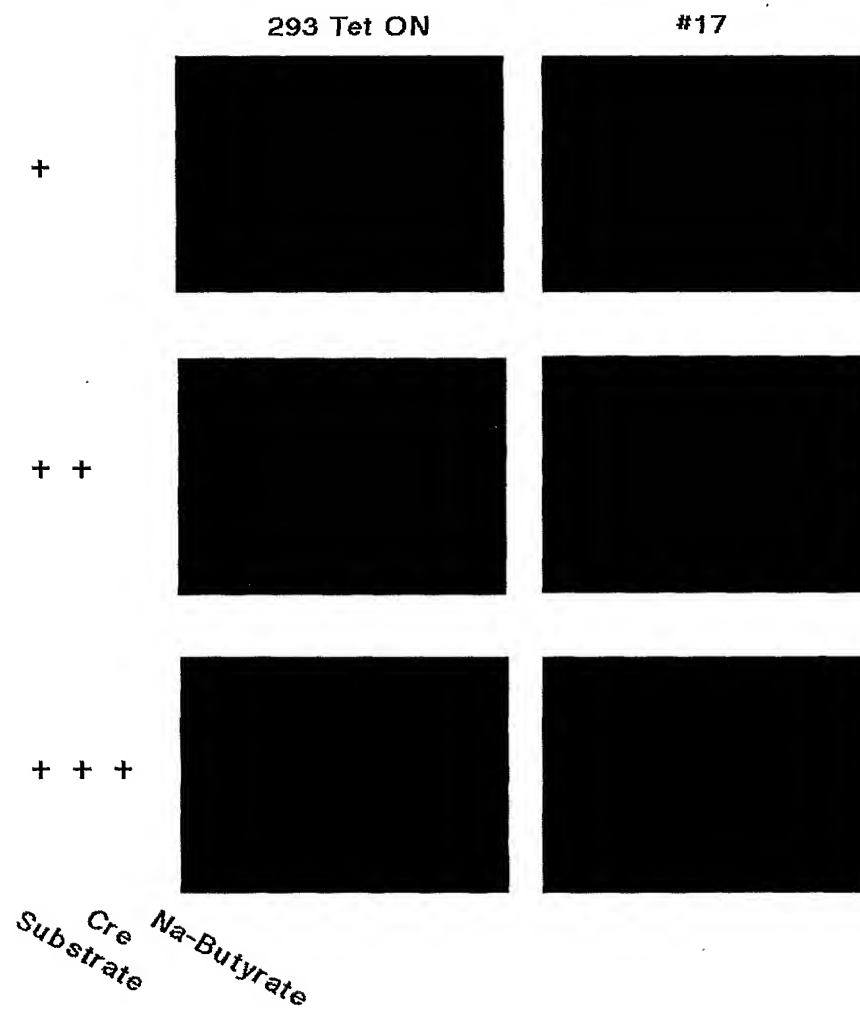
24/40

FIG. 10E



25/40

FIG. 11



26/40

FIG. 12

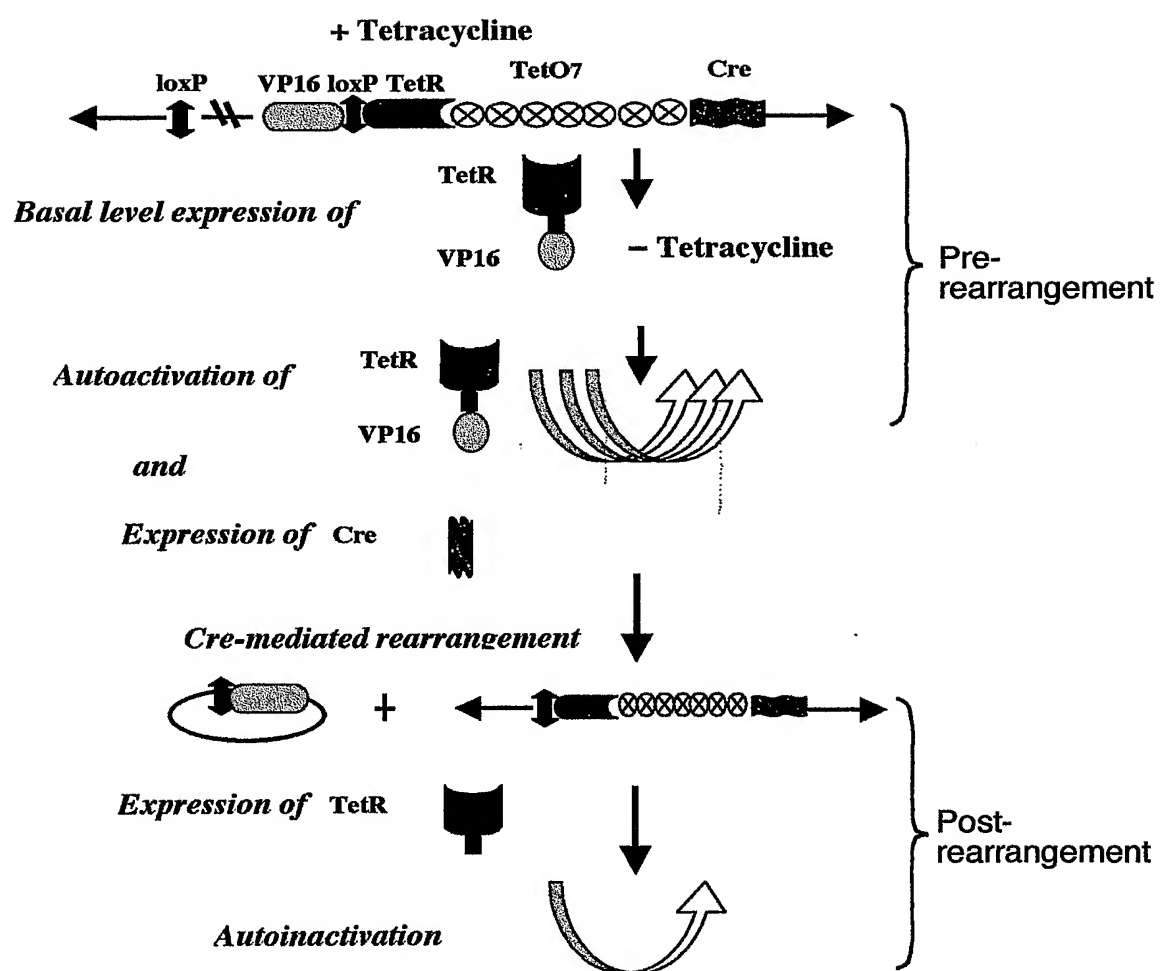
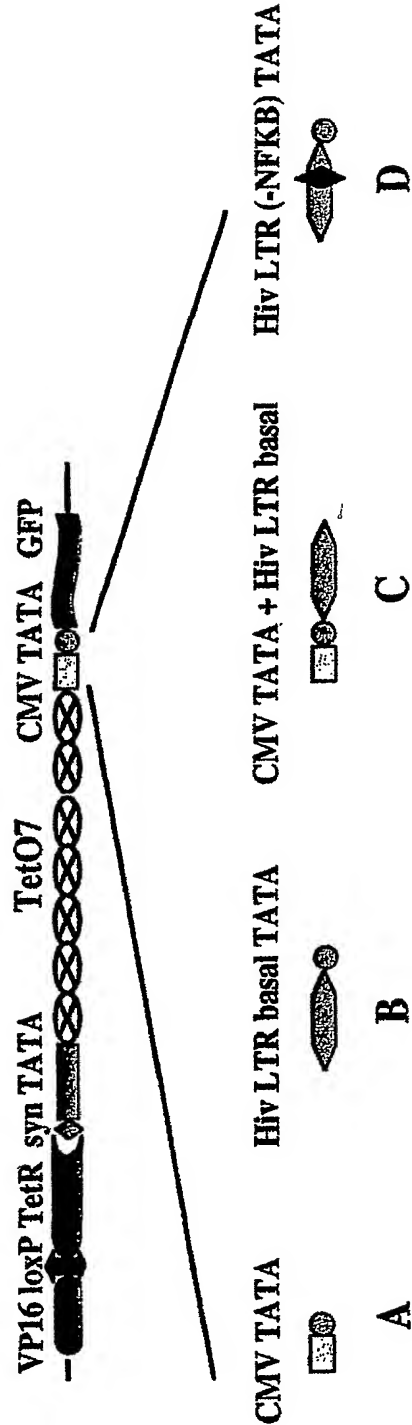


FIG. 13A



28/40

FIG. 13B

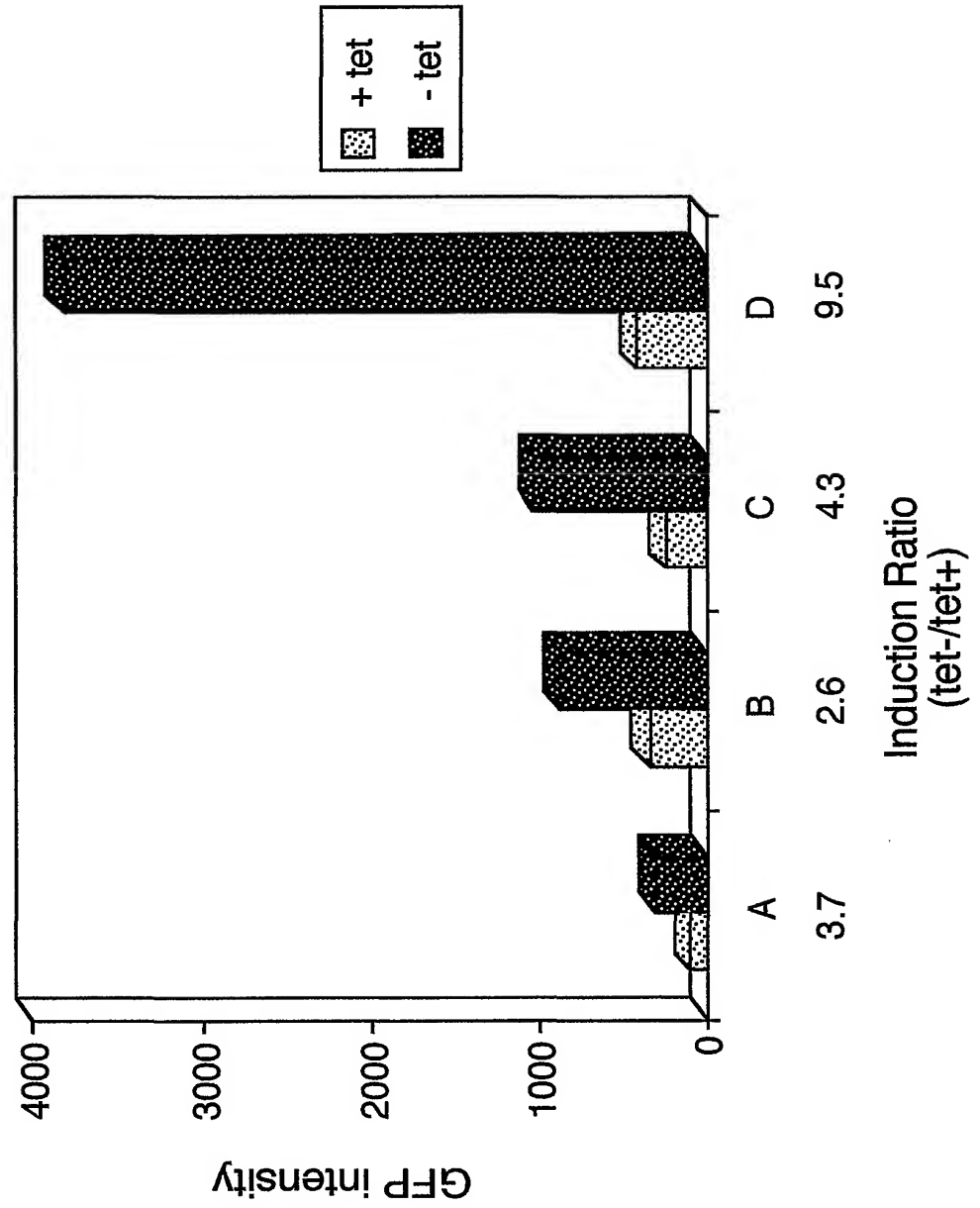
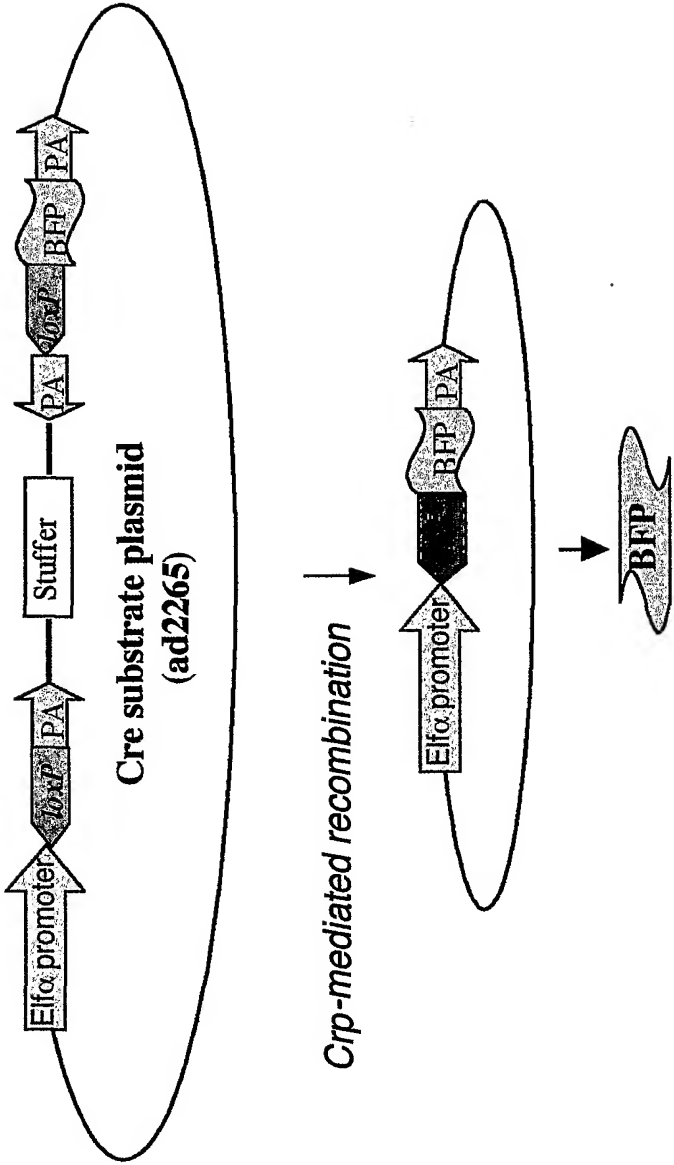


FIG. 14



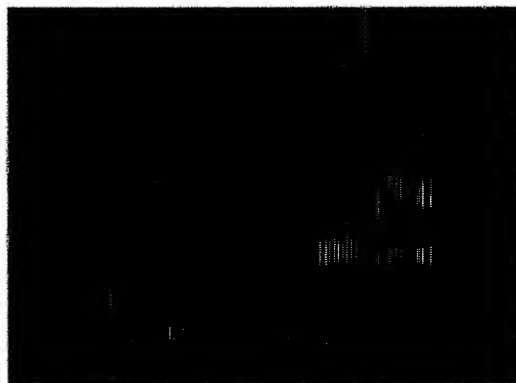
30/40

FIG. 15A

0 μ M Est.

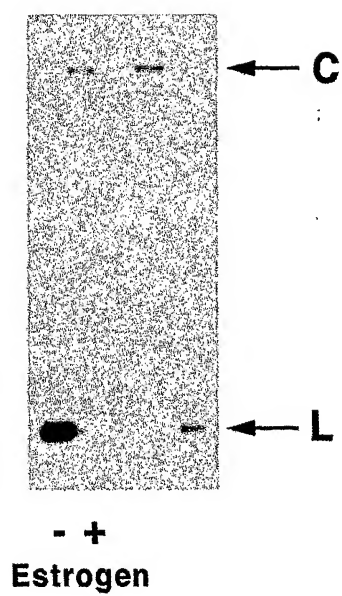


2 μ M Est.



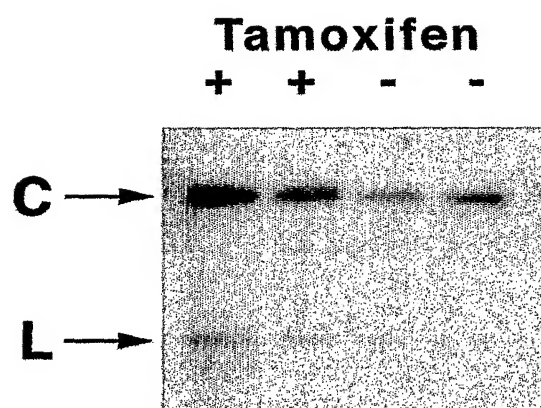
31/40

FIG. 15B



32/40

FIG. 16



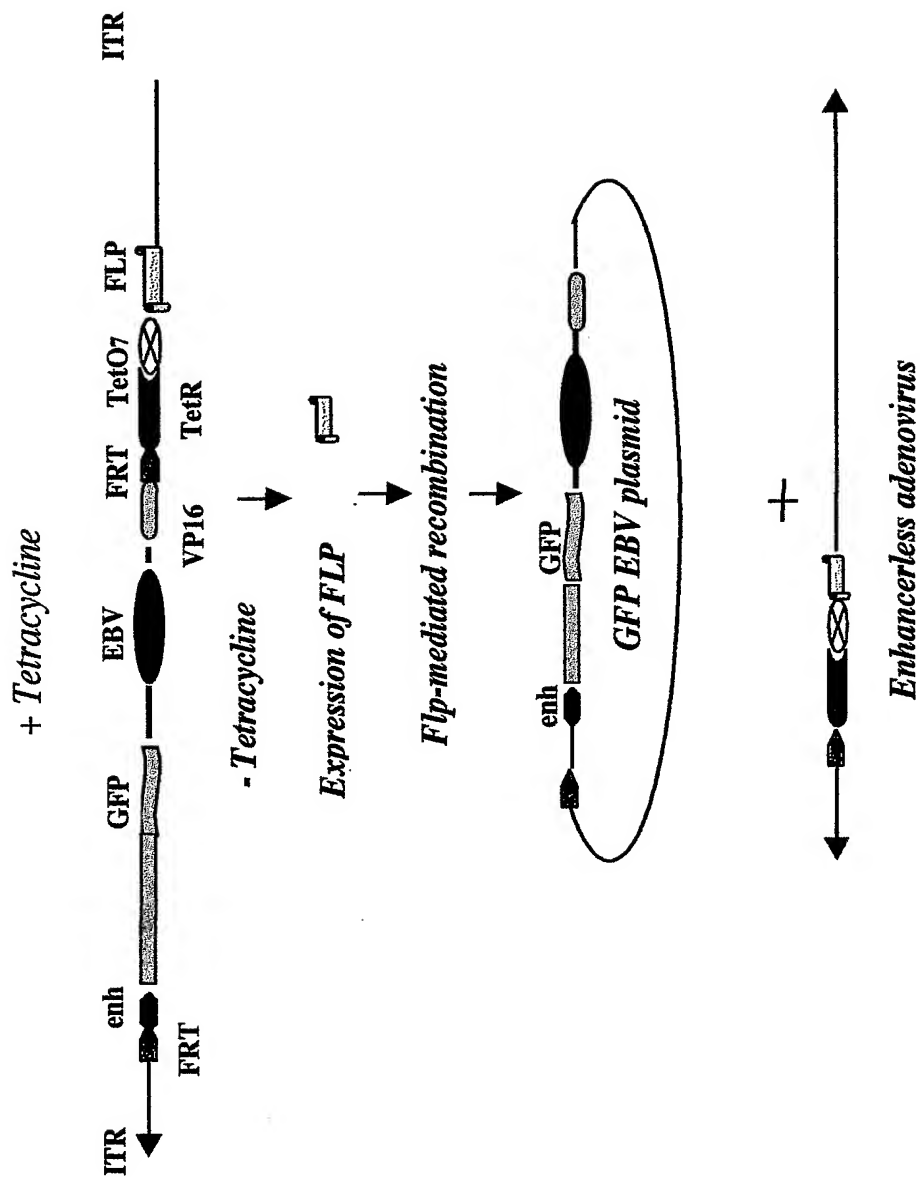
33/40

FIG. 17



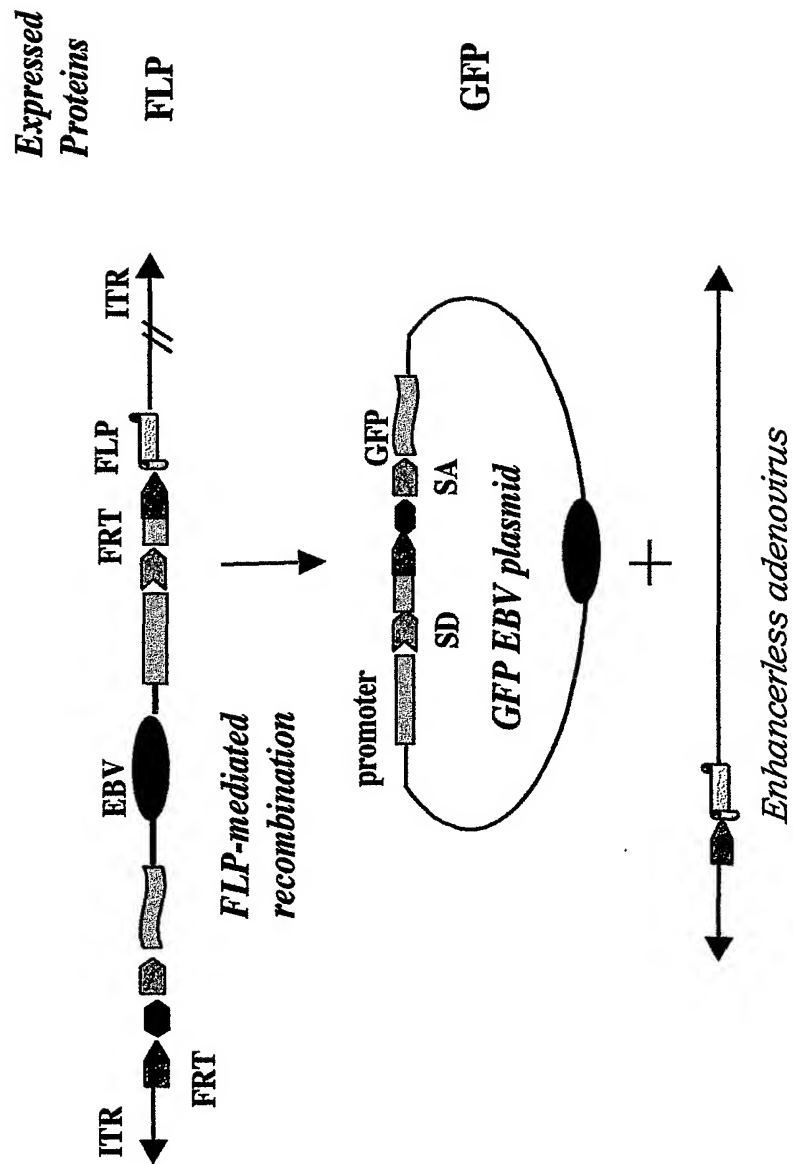
34/40

FIG. 18A



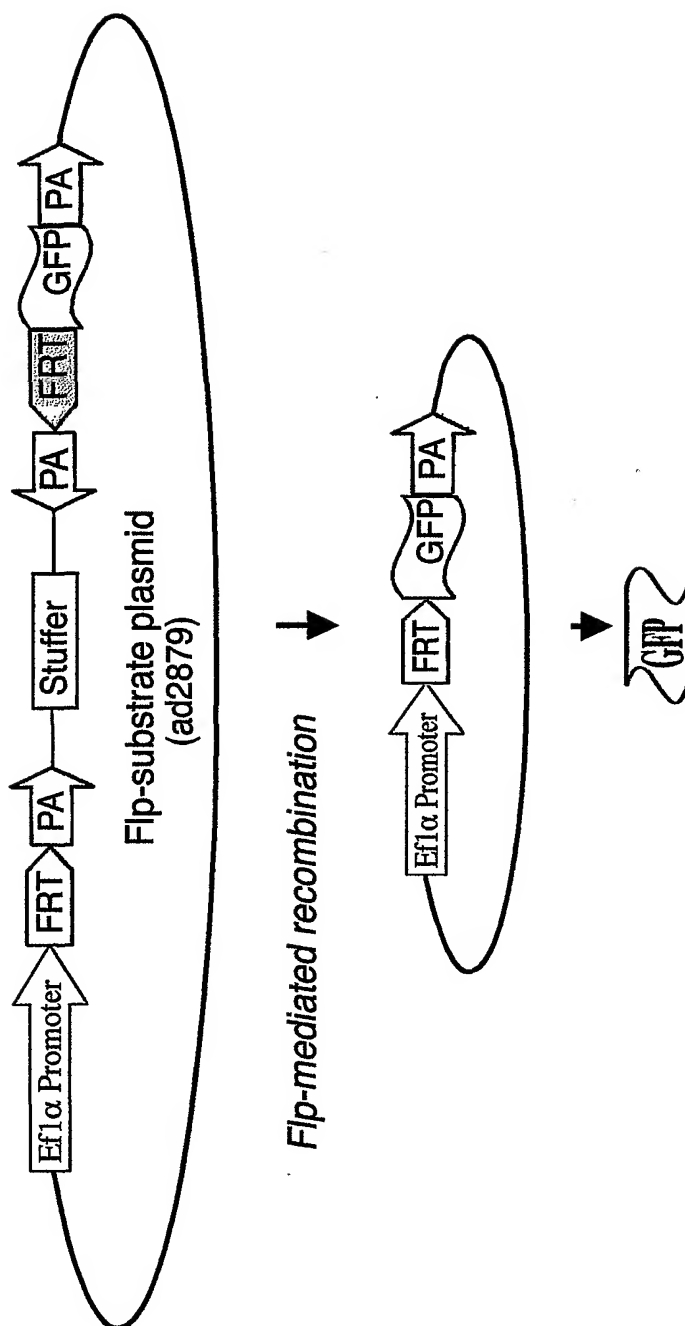
35/40

FIG. 18B



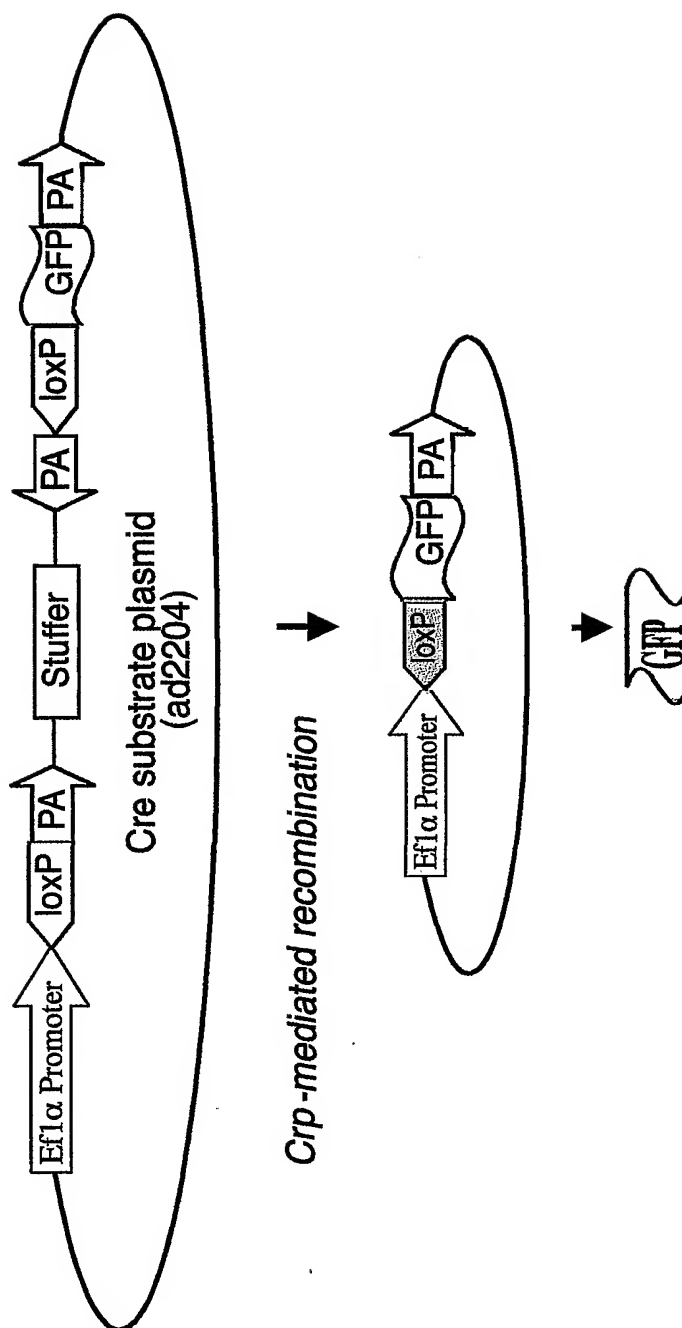
36/40

FIG. 19A



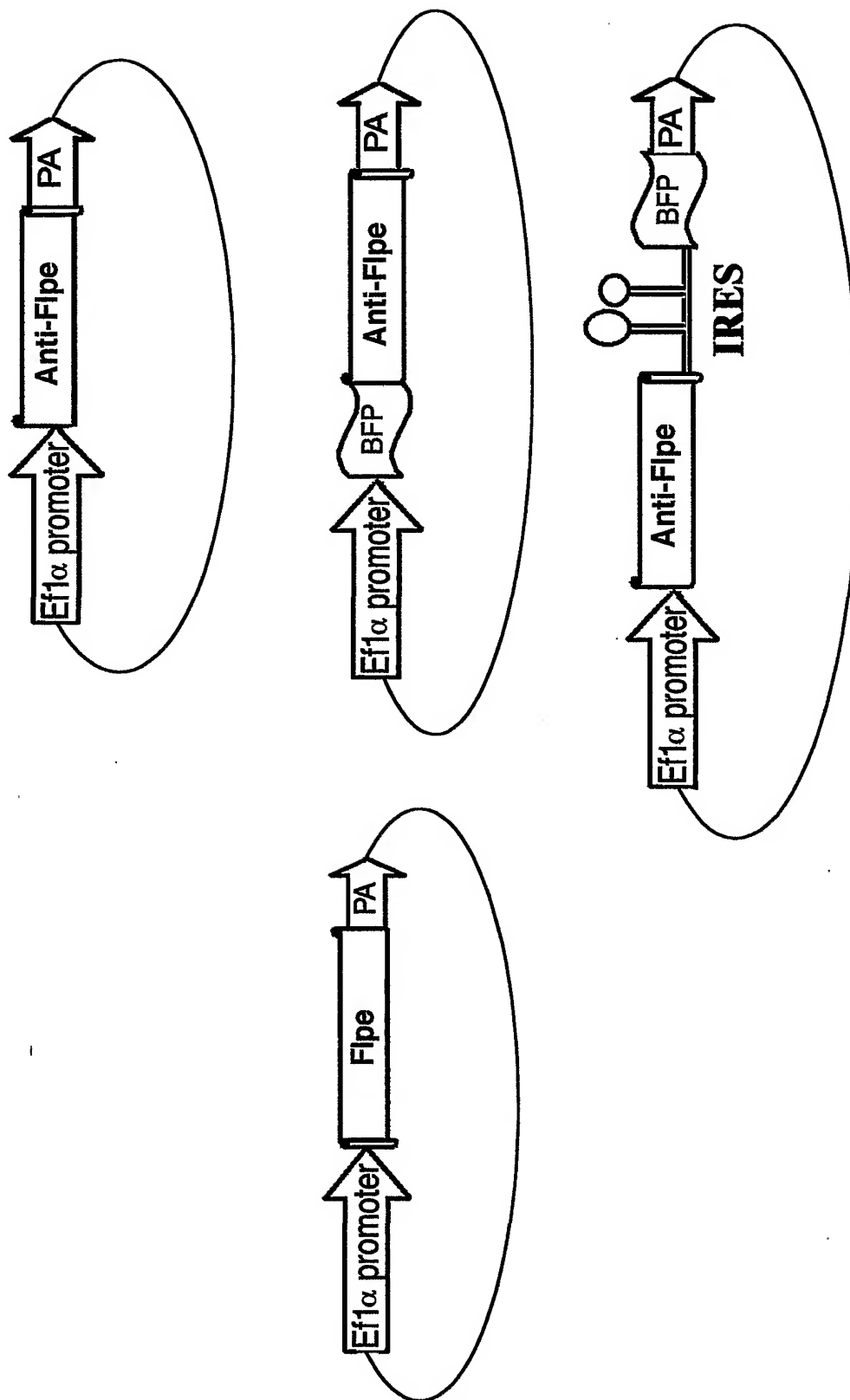
37/40

FIG. 19B



38/40

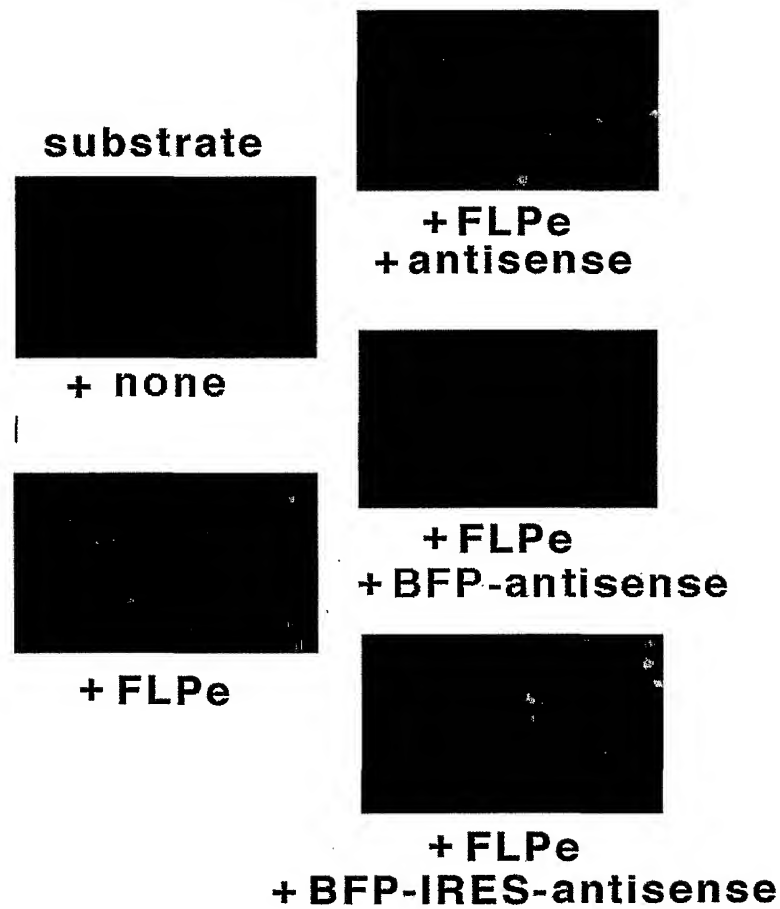
FIG. 20





39/40

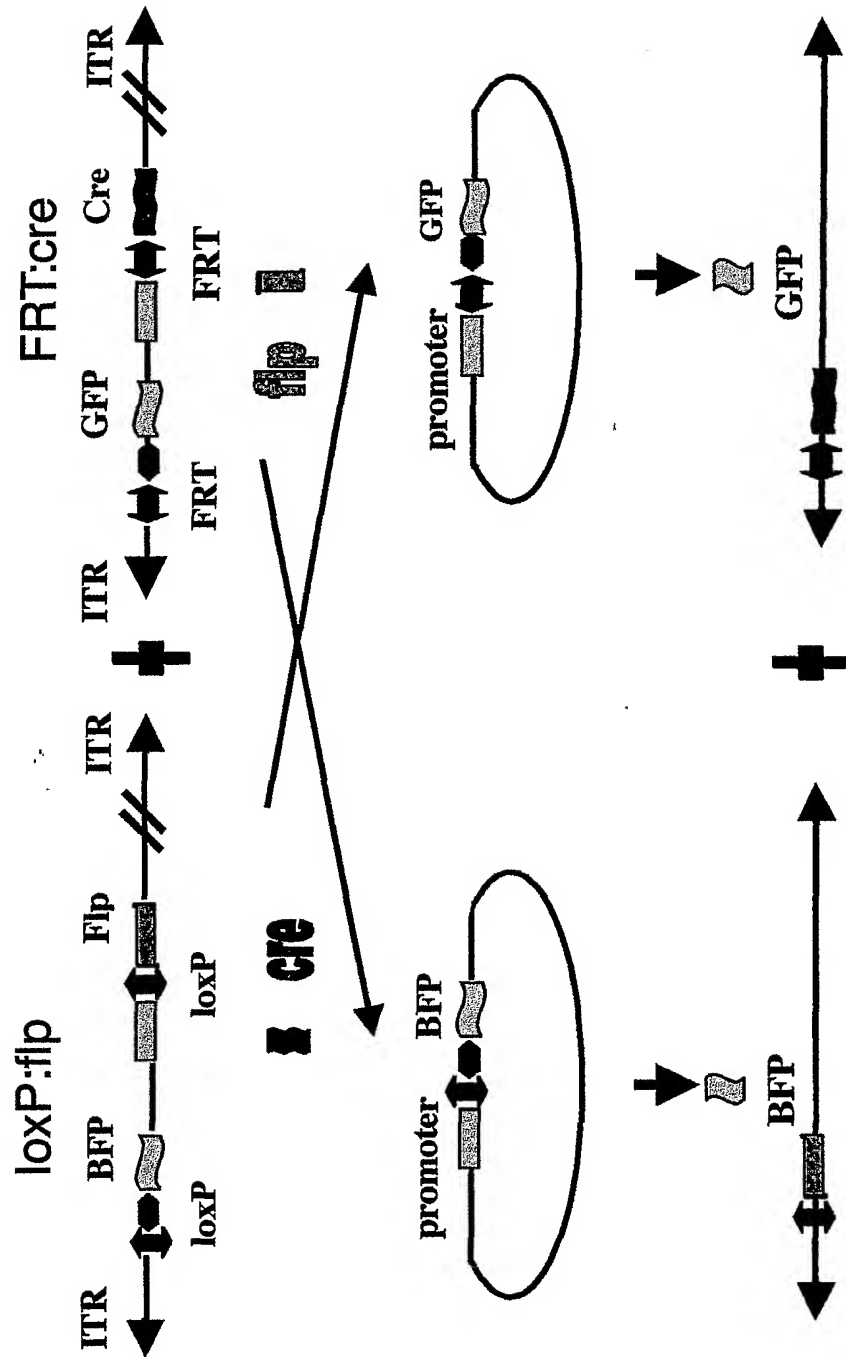
FIG. 21



RO/L

40/40

FIG. 22



SEQUENCE LISTING

<110> The General Hospital Corporation

<120> Self-rearranging DNA vectors

<130> 00786/352WO3

<150> US 60/231,053

<151> 2000-09-08

<150> US 60/246,904

<151> 2000-11-08

<160> 10

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 2341

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 1

```

gaatcggccca gcgcgaattc gattatcatc atcataatat accttatttt ggattgaagc 60
caatatgata atgaggggggt ggagtttgtg acgtggcgcg gggcgtagga acggggcggg 120
tgacgtaggt tttagggcgg agtaacttgc atgtattggg aattgtagtt tttttaaaat 180
gggaagttac gtacgcggca tcgatgcgcg ggatatcgcg gcggctagcg acatgagggt 240
gccccgtatt cagtgtcgct gatttgtatt gtctgaagtt gtttttacgt taagttgatg 300
cagatcaatt aatacgatac ctgcgtcata attgattatt tgacgtggtt tgatggcctc 360
cacgcacgtt gtgatatgta gatgataatc attatcaactt tacgggtcct ttccggtgat 420
ccgacagggt acggggcggc gacctcgcg gttttcgcta tttatgaaaa tttccgggtt 480
taaggcggtt cgttcttctc tcgtcataac ttaatgtttt tatttaaaat accctctgaa 540
aagaaaggaa acgacagggt ctgaaagcga ggcttttttg cctctgtcgt ttcttttctc 600
tgtttttgtc cgtggaatga acaatggaag ttaacggatc caggccgcga gcaaaaggcc 660
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 720
cccctgacga gcatcacaaa aatcaacgct caagtcagag gtggcgaaac ccgacaggac 780
tataaagata ccaggcggtt ccccttgaa gctccctcgt gcgctctcct gttccgacct 840
tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 900
gctcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc 960
acgaaccccc cgttcagccc gaccgctgcy ccttatccgg taactatcgt cttgagtcca 1020
acccgtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 1080
cgagggtatg aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 1140
gaagaacagt atttggtatc tgcgctctgc caaagccagt taccttcgga aaaagagttg 1200
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 1260
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 1320
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatcaga ttatcaaaaa 1380
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 1440
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 1500
tctgtctatt tcgttcaccc atagttgcct gactccccgt agtgtagata actacgatac 1560
gggagggctt accatccggc cccagtgtcg caatgatacc gcgtgaccca cgctcaccgg 1620
ctcctgattt atcagcaata aaccagccag ccggaagtgc cgagcgcaga agtggtcctg 1680
caactttatc cgcctccatc cagtctatta gttgttgccg ggaagctaga gtaagtagtt 1740

```

```

cgccagttaa tagtttttcgc aacgtttgttg ccattgctac aggcacgtg gtgtcacgct 1800
cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 1860
cccccatgtt gtgcaaaaaa gcggttagct ccttcgggtcc tccgatagtt gtcagaagta 1920
agttggccgc agtggttatca ctcatgggta tggcagcact gcataattct cttactgtca 1980
tgccatccgt aagatgcttt tctgtgactg gtgagtattc aaccaagaat acgggataat 2040
accgcgccac atagcagaac tttaaaagtg ctcatcattg ggaaacgttc ttcggggcga 2100
aaactctcaa ggatcttacc gctgttgaga tccagttcga tgaacccac tcgcgcaccc 2160
aagtgatctt ctgcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg 2220
caaatgccc caaaaaaggg aataagggcg acacggaaat gttgaatact catacttttc 2280
ctttttcaat attattgaag catttatcag ggttattgtc tcatcagcgg atacatattt 2340
g

```

<210> 2

<211> 34616

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 2

```

gaatcggcca ggcggaatta actataacgg tcctaaggta gcgtcatcat cataatatac 60
cttatttttg attgaagcca atatgataat gagggggtgg agtttgtgac gtggcgcggg 120
gcgtgggaac ggggcgggtg acgtagggtt tagggcggag taacttgcac gtattgggaa 180
ttgtagtttt tttaaaatgg gaagttacgt atcgtgggaa aacggaagtg aagatttgag 240
gaagtgtgtg gtttttttggc tttcgtttct gggcgtaggt tcgcgtgcgg ttttctgggt 300
gttttttttg gactttaacc gttacgtcat ttttttagtcc tatatatact cgtctgttac 360
ttggcccttt ttacactgtg actgattgag ctggtgccgt gtcgagtggg gttttttaat 420
aggttttttt actggtaagg ctgactgtta tggctgccgc tgtggaagcg ctgtatgttg 480
ttctggagcg ggagggtgct attttgccta ggcaggaggg tttttcaggt gtttatgtgt 540
ttttctctcc tattaatttt gttataacct ctatgggggc tgtaatgttg tctctacgcc 600
tgccgggtatg tattcccccg ggctatttct gtcgcttttt agcactgaac gatgttaacc 660
aacctgatgt gtttaccgag tcttacatta tgactccgga catgaccgag gaactgtcgg 720
tggtgctttt taatcacggg gaccagtttt tttacgggtc cgcggcatg gccgtagtcc 780
gtcttatgct tataagggtt gtttttcttg ttgtaagaca ggcttctaata gtttaaattgt 840
tttttttttg tatttttatt tgtgtttaat gcaggaaacc gcagacatgt ttgagagaaa 900
aatgggtgtc ttttctgttg tggttccgga acttacctgc ctttatctgc atgagcatga 960
ctacgatgtg cttgcttttt tgccgcaggc tttgcctgat tttttgagca gcaccttgca 1020
ttttatatcg ccgcccatgc aacaagctta cataggggct acgctgggta gcatagctcc 1080
gagatgtggt gtcataatca gtgtgggttc ttttgtcatg gttcctggcg ggtgaaggc 1140
cgcgctggtc cgtgcagacc tgcacgatta tgttcagctg gccctgcgaa gggacctacg 1200
ggatcgcggt atttttgtta atgttccgct tttgaatctt atacaggtct gtgaggaacc 1260
tgaatttttg caatcatgat tcgctgcttg aggctgaagg tggagggcgc tctggagcag 1320
atttttacaa tggccgggact taatatccgg gatttgctta gagacatatt gataaggttg 1380
cgagatgaaa attatttggg catggttgaa ggtgctggaa tgtttataga ggagattcac 1440
cctgaagggt ttagccttta cgtccacttg gacgtgaggg cagtttgctt tttggaagcc 1500
attgtgcaac atcttacaaa tgccattatc tgttcttttg ctgtagagtt tgaccacgcc 1560
accggagggg agcgcgttca cttaatagat cttcattttg aggtttttgga taatcttttg 1620
gaataaaaaa aaaaaaaca tggttcttcc agctcttccc gctcctcccg tgtgtgactc 1680
gcagaacgaa tgtgtagggt ggctgggtgt ggcttattct gcggtggttg atgttatcag 1740
ggcagcggcg catgaaggag tttacataga acccgaagcc agggggcgcc tggatgcttt 1800
gagagagtgg atatactaca actactacac agagcgagct aagcgacgag accggagacg 1860
cagatctgtt tgtcacgccc gcacctggtt ttgcttcagg aaatatgact acgtccggcg 1920
ttccatttgg catgacacta cgaccaacac gatctcgggt gtctcggcgc actccgtaca 1980
gtagggatcg cctacctctt tttgagacag agaccgcgcg taccatactg gaggatcatc 2040
cgctgctgcc cgaatgtaac actttgacaa tgcacaacgt gagttacgtg cgaggtcttc 2100
cctgcagtgt gggatttacg ctgattcagg aatgggttgt tccctgggat atggttctga 2160
cgccgggagga gcttgtaatc ctgaggaagt gtatgcacgt gtgcctgtgt tgtgccaaca 2220

```

ttgatatcat	gaogagcatg	atgatccatg	gttacgagtc	ctgggctctc	cactgtcatt	2280
gttcaggtcc	cggttccctg	cagtgcatag	ccggcgggca	ggttttggcc	agctgggtta	2340
ggatgggtgg	ggatggcgcc	atgtttaatc	agaggtttat	atggtaaccg	gaggtgggtga	2400
attacaacat	gccaaaagag	gtaatgttta	tgtccagcgt	gtttatgagg	ggtcgccact	2460
taatctacct	gcgcttggtg	tatgatggcc	acgtgggttc	tgtgggtccc	gccatgagct	2520
ttggatacag	cgccttgcat	tgtgggattt	tgaacaatat	tgtgggtgctg	tgctgcagtt	2580
actgtgctga	tttaagttag	atcaggggtg	gctgctgtgc	ccggaggaca	agggcgtctca	2640
tgctgcgggc	ggtgcgaatc	atcgctgagg	agaccactgc	catgttgtat	tcctgcagga	2700
cggagcggcg	gcggcagcag	tttattcgcg	cgctgctgca	gcaccaccgc	cctatcctga	2760
tgcacgatta	tgactctacc	cccatgtagg	cgtggacttc	cccttcgcgc	ccggttgagc	2820
aaccgcaagt	tggaacagcag	cctgtggctc	agcagctgga	cagcgacatg	aacttaagcg	2880
agctgcccgg	ggagtttatt	aatatcactg	atgagcgttt	ggctcgacag	gaaaccgtgt	2940
ggaatataac	acctaagaat	atgtctgtta	cccatgatat	gatgcttttt	aaggccagcc	3000
ggggagaaag	gactgtgtac	tctgtgtgtt	gggagggagg	tggcagggtt	aatactaggg	3060
ttctgtgagt	ttgattaaag	tacggtgatc	aatataagct	atgtgggtgg	ggggctatac	3120
tactgaatga	aaaatgactt	gaaattttct	gcaattgaaa	aataaacacg	ttgaaacata	3180
acatgcaaca	ggttcacgat	tctttattcc	tgggcaatgt	aggagaaggt	gtaagagttg	3240
gtagcaaaag	tttcagtggt	gtattttcca	ctttccagg	accatgtaaa	agacatagag	3300
taagtgttta	cctcgctagt	ttctgtggat	tcactagtgc	cattaagtgt	aatggtaagt	3360
atcatagggt	tagttttatc	accatgcaag	taaacttgac	tgacaatgtt	attttttagca	3420
gtttgacttt	gggttttttg	ataggctaga	aggttaggca	taaatccaac	tgcatthtgt	3480
tatggatttc	cattagttag	gttcccattt	ctaaagtthc	agtaatgttt	tttaagttag	3540
gagttctcca	ttagaacacc	gttttgggtc	aatctaagga	atatactaac	acttgcaacg	3600
gtgcctgtca	tggatgaaag	atctccagat	acagccaaag	cagctacagt	agctagtact	3660
tgactcccac	atthttgtaag	aaccaaagta	aattttgcagt	cattatctga	atgaattctg	3720
cagttaggag	atgggtcttg	ggttgtccac	agggtaagtt	tgtcatcatt	tttgtttcct	3780
attgtaatgg	cccctgagtt	gtcaaagctt	aaaccgcgtc	caagtttagt	aatcatggca	3840
ccgttttcat	tgtaatcaat	gccagagcca	atthttagtt	ttattgggtt	gatattctgga	3900
gactcctcca	tgtttgtatc	aaactccaga	ccctttcctg	catttatagc	tatggcagta	3960
ttatcaaagt	ttagtccact	ggattttttt	atgctaactt	ccagtttttt	agtattgttt	4020
gatgcattaa	aaaggtatag	gcctctgtta	tagtttatgt	ccaagttatg	agatgcatta	4080
atatacaggg	gtccctgccc	cagtttaaga	cgtagttttg	tttgagcatc	aaatgggtaa	4140
tcacacatcta	gaattaacaa	gttgttattt	atacgcatgc	caccgcccgt	tttaattttcc	4200
atgthgtttg	atgaatcata	accaatagct	cctgcaactt	tggttctaag	ggagthtttg	4260
tcaacgggtga	cacctgggtcc	agtaactact	atctagtgtat	cggagthttg	tgctacttgc	4320
aaaggaccgc	ttatthttaat	tcctattttt	ccattattta	cataaatagg	atcttccatg	4380
thaatgccc	agctaccctg	ggcagtagtt	agcgggggtg	atgcagttac	agtaaggggtg	4440
tcgctgtcac	tgccagagag	gggggctgat	gtttgcaggg	ctagctttcc	atctgacact	4500
gtaatggggc	ctthtagtagc	aatgcttagt	ttggagtctt	gcacgggtcag	tggggcttgt	4560
gactgtacgc	taagagcgcc	gctagtaact	atcagaggag	cgttggttgc	cactgttttg	4620
gcgctgagg	taattgtaag	tgggtcggag	gtgtccaaac	ttatgtttga	ctthgttttt	4680
thaaagtggct	gagtaacagt	ggttacattt	tgggaggtga	ggthttccgg	cttgtctagg	4740
gtaagaccgc	tgcccattht	aagcgcaagc	atgccgtggg	agggtgtccaa	aggthtcggag	4800
acgcgtagag	agagaactcc	agggggactt	tcttggaac	cattgggtga	aacaaatgga	4860
ggggtaagaa	agggcacagt	tggaggcccc	gtthctgtgt	catatggata	cacgggggtt	4920
aaggthtctt	cagacggtct	ggcgcgthtt	atctgcaaca	atatgaagat	agthgggtgc	4980
gagggacaag	aacatgagga	atthgacatc	ccattthaaac	thtgagaaaa	gtthtgagct	5040
aaaaggcggc	tgagatacca	gagthgggag	gaaggaaaag	aggthgatgt	gaataagctg	5100
gacaaagatt	tgctgactga	ththtaagtaa	gtaattthatt	gtgtgtthtat	gttagthtgaa	5160
tgggaataaga	tctctaatac	cacacatggt	ththaaataga	gtgcagaggt	cctctggacc	5220
ctgatagggg	aagtgcaggc	agccctctgt	thctgccgag	tgctgggtga	cggthgatagg	5280
ththttctccc	accataagca	ccagththtt	gcgctgggtg	gthtagcttht	agctgaggcg	5340
gttgccggta	gtggththtt	cgtaggtaag	ththggcctgc	thgaccacac	aaaagatacc	5400
tctthttacac	tggtgttaggt	taaccatgtc	thcaacttht	thththtaggc	gtthctcgctc	5460
ggacgcggcc	thgcccctth	ctagtagggc	ctgttcgggtg	thaatthccat	ccaattctag	5520
atctagagat	tcagthcatct	ccacctgtca	aathaaagta	gthaatctca	gtgggggtgg	5580
gagaaggggg	gcgaggctga	thgattgggg	caataacctg	thgcagthgt	atgacagcgg	5640
gcactgggaa	agtagggtgg	thcatggcat	ctatggcatt	ccagccaatg	thaaaggtatg	5700

gatatatggc	tagggcaaaa	atgggtactgc	aaaaaaccat	gacagagatg	atggcgtata	5760
accaggcttc	tgacaaatcg	ctctgtttgt	tgtagcagct	gggaatgttc	catatttgag	5820
tgaatctgca	ggaaatatgt	cttttgggag	gcgctgaggt	ttgggagcaa	agcacaggta	5880
gggcgcaaaa	aatcagcaaa	acaaaaatga	cactccgttt	cataatataa	gaattctgag	5940
aagatcagct	atagtccgtg	ctctgtattg	cggatgggtgc	ctgagggtacg	caatgcgcac	6000
acaaaccag	tcaatgaact	gaatgaaggc	gatgactaca	gtgacgaggc	tgcagatgag	6060
gataaggggtg	acaaatccgt	aaagcaggta	aactgtgaaa	ggtgggatgc	aatctacttc	6120
gatgtgagcg	accgcggcca	atgtagagca	cgcacagaaa	agcgcacaac	gggtcaataa	6180
tataagaact	cgaggaatca	tgtctcattt	aatcatactg	taaaagaaga	gaacatgggt	6240
tcagaccgtc	caatctatga	attttttcat	tgtgtgggtt	gagcacaatg	ataggcctat	6300
agatgggggg	tctggcgcgt	ctgcgcttta	ggcaacaaat	aagccacata	ataataaggc	6360
aaacaaacat	aagcgctatg	gaaaaccacc	acatgtccaa	gctcgcccag	tcattgacaa	6420
aggcatgaac	ttggggtaaa	tttagggcag	atgttagtcc	ggtagcagtg	gtgttgcgat	6480
agtccgttgt	gggcgcgatg	gttgagccgg	tcattctctgg	agcaggcaag	ctgaagctgg	6540
gtttgatcaa	atttgcagtg	caggcgctgg	cagaaatcag	gcgctaactg	ccaggaaagt	6600
ttgatttgaa	ggttgtgggt	ataatctgg	cgcgctggag	catatcccac	atagagataa	6660
ttgtccaggg	gaatacaagc	aagcggaaaa	tcaaggcatt	ttcttttcat	caataaaaact	6720
gcgtctgctt	ttgtatttga	gataaagtaa	ggtacatacc	aaagcaagcg	ctgtaataag	6780
cagagcgggtg	gaacaaaagg	tgccagtgtt	ctctaaacac	ttttgtgggg	gccacaactt	6840
gtactgtttg	ctcatgtaca	tggtaatatc	gcacatttca	taaaatggaa	atttatacat	6900
aaaagtttta	cgattttcac	cttggaagac	tgtgacatta	tagtcgttag	tgtcacctgg	6960
ctgccaaata	gcatatacag	catacttgcc	aatttttgtct	ttgtggcgaa	taataagctt	7020
ttcatgttct	gtggtgcatt	ttataagagt	agtgcatcca	ttagcttctg	atttaaatgt	7080
aacattgcaa	gctggttcct	taaactcaac	ctttttggca	gcgctgcaga	ctgccgcaag	7140
ggcgagcaag	cctaaaatca	tgtacctcat	cttggtatgt	gccccagcg	tttaaaaagc	7200
tgacaatagg	tacaaacgtg	cgtgcagcag	gcggcaaccc	taaggcacag	aagtgtctag	7260
ataagaataa	acagaattac	aagagtaagg	ataaccccga	ccccaatcc	agaaaaatta	7320
gacaagcttg	tagagttact	tgaattgctc	atatacttaa	ttaaaaaatc	ccagcacccc	7380
gcaaaatgct	tttttgacct	gagttccggg	agttgagctc	acctcctgtt	ttggaaaaat	7440
gggagtaatg	tctggttacg	ctcaggctgt	agggtgtggg	gcagcaaccg	gtgacgcact	7500
cgtacgttcc	cggcaggtga	ggaggggtgt	ggtggtggtg	tttttcttga	cggtgtagtt	7560
gaagccgaga	aggttgtgtg	gcaaacctac	ttcgtctcgc	tggaaactgt	tgtaaattac	7620
aatgaagag	ccgttaaagt	accaggtaag	gtacttattg	gcccgttgt	gcaaaccgga	7680
ggtgagggtt	gctttggtct	gctttgggtg	gtgtaaaaac	gtggcgttca	caggatggcg	7740
acaggagccc	cagtagattc	taatttctgt	atattattata	ctcagcacag	agatgacaac	7800
aaagatcttg	atgtaatcca	gggttaggac	agttgcaaac	cacggtcaga	acacagggac	7860
cccgtcccg	ctccactagc	agggggcgct	tggtaaactc	ccgaatcagg	ctacgtgtaa	7920
gctctacctg	ggtggtgagc	cggacgccgt	gcgccggggc	ctcgatatgc	tcttcggggc	7980
attcaaagta	acaaaactca	ccggagccgc	gggcaaagca	cttgtggcgg	cggcagtggt	8040
cgagggtgtg	caggcgagtg	cgctctgcct	ctccactggg	cattcagtcg	taccgctccg	8100
ccgagttctt	caccgcgtca	aagttgggaa	taaactggtc	cgggtagtg	ccgggaggtc	8160
cagaaaaggg	gttgaagtaa	accgaaggca	cgaactcctc	aataaattgt	agagttccaa	8220
tgctcccgga	gcgcgggtcc	gaggacgagg	tctgcagagt	taggatcgcc	tgacggggcg	8280
taaatgaaga	gcggccagcg	ccgccgatct	gaaatgtccc	gtccggacgg	agaccaagag	8340
aggagctcac	cgactcgtcg	ttgagctgaa	tacctcgccc	tctgattttc	agggtagtta	8400
tacctgccc	ggcgaccgc	accctgtgac	gaaagccgcc	cgcaagctgc	gcccctgagt	8460
tagtcatctg	aacttcggcc	tgggcgtctc	tgggaagtac	cacagtgggtg	ggagcggggc	8520
tttcttggtg	caccagggca	gcgggccaac	tacgggggatt	aaggttatta	cgagggtgtg	8580
tggtaatagc	cgcctgttcg	aggagaattc	ggtttcgggtg	ggcgcggtat	ccgttgacct	8640
gggatatacat	gtgggggtccc	gcgctcatgt	agtttattcg	ggttgagtag	tcttggggcag	8700
ctccagccgc	aagtcctcat	tgtggctggt	aactccacat	gtagggcgtg	ggaatttctt	8760
tgctcataat	ggcgctgacg	acagggtgct	gcgcgggtg	tggccgctgg	agatgacgta	8820
gttttgcgcg	ttaaatttga	gaaagggcgc	gaaactagtc	cttaagagtc	agcgcgcagt	8880
atlttgctgaa	gagagcctcc	gcgtcttcca	gcgtgcgccg	aagctgatct	tcgcttttgt	8940
gatacaggca	gctgcgggtg	agggagcgca	gagacctgtt	ttttattttc	agctcttggt	9000
cttggccccc	gctttgttga	aatatagcat	acagagtggg	aaaaatccta	tttctaagct	9060
cgcgggtcga	tacgggttcg	ttgggcgcga	gacgcagcgc	tctcctcctt	gctgctgccg	9120
ccgctgtgga	tttcttgggc	tttgtcagag	tcttgcctac	cggtcgcctt	tgcttctgtg	9180

tgaccgctgc	tgttgctgcc	gctgccgctg	cgcgcggtgc	agtaggggct	gtagagatga	9240
cggtagtaat	gcaggatggt	acgggggaag	gccacgcctg	gatggtagag	aagaaagcgg	9300
cgggcgaagg	agatggtgcc	cccacagtct	tgcaagcaag	caactatggc	gttcttgtgc	9360
ccgcgccacg	agcggtagcc	ttggcgctgt	tgttgctctt	gggctaaccg	cggcggctgc	9420
ttagacttac	cggccctggt	tccagtgggt	tcccatctac	ggttgggtcg	gcgaacaggc	9480
agtgcgcgcg	gcgcctgagg	agcggagggt	gtagcgatgc	tgggaacggt	tgccaatttc	9540
tggggcgccg	gcgaggggaa	tgcgaccgag	ggtgacggtg	tttcgtctga	cacctcttcg	9600
gcctcggaag	cttcgtctag	gctgtcccag	tcttccatca	tctcctcctc	ctcgtccaaa	9660
acctcctctg	cctgactgtc	ccagtattcc	tcctcgcccg	tgggtggcgg	cggcggcagc	9720
tcgagcttct	ttttgggtgc	catcctggga	agcaagggcc	cgcggctgct	gatagggctg	9780
cggcggcggg	gggatgggtg	tgagctcctc	gccggactgg	gggtccagggt	aaaccccccg	9840
tccctttcgt	agcagaaact	cttggcgggc	tttggtgatg	gcttgcaatt	ggccaaggat	9900
gtggcccttg	gtaatgacgc	aggcggtaag	ctccgcattt	ggcgggcggg	attggtcttc	9960
gtagaaccta	atctcgtggg	cgtggtagtc	ctcagggtaca	aatttgcgaa	ggtaagccga	10020
cgtccacagc	cccggagtga	gtttcaaccc	cggagcccg	gacttttcgt	caggcgaggg	10080
accctgcagc	tcaaagggtac	cgataatttg	acttttcgta	agcagttgcg	aattgcagac	10140
cagggagcgg	tgccgggtgc	ataggttgca	ggcagcagta	cactccagta	ggccctcacc	10200
gctcacgtct	tccatgatgt	cggagtggta	ggcaaggtag	ttggctagct	gcagaaggta	10260
gcagtgacct	caaagcggcg	gagggcattc	acggctactta	atgggacaaa	agtcgctagg	10320
aagcgcacag	caggtggcgg	gcagaattcc	tgaacgctct	aggataaagt	tcctaaagtt	10380
ttgcaacatg	ctttgactgg	tgaagtctgg	cagaccctgt	tgcagggttt	taagcaggcg	10440
ttcgggggaag	ataatgtccg	ccaggtgcgc	ggccacggag	cgctcgttga	aggccgtcca	10500
taggtccttc	aagttttgct	ttagcagctt	ctgcagctcc	tttaggttgc	gtccctccag	10560
gcattgctgc	cacacgcca	tggcgtttg	ccagggttag	cacagaaata	agtaaagca	10620
gtcgcggacg	tagtcgcggc	gcgcctcgcc	cttgagcgtg	gaatgaagca	cgttttgccc	10680
gagggcggtt	tcttgcaaaa	ttccaaggta	ggagaccagg	ttgcagagct	ccacgttgga	10740
aattttgcag	gcctggcgca	cgtagccctg	gcgaaagggt	tagtgcaacg	tttcctctag	10800
cttgcgctgc	atctccgggt	cagcaaagaa	ccgctgcacg	cactcaagct	ccacggtaac	10860
aagcacctgc	gccaatatta	gcttgctcgc	ctcctccaag	tcggcagggt	cgcgcgtctc	10920
aagccagcgc	gccagctgct	catcgccaac	tgccgggtagg	ccctcctcgg	tttgttcttg	10980
caagtttgca	tccctctcca	ggggctcgtg	acggcgcaacg	atcagctcgc	tcatgactgt	11040
gctcataaac	ttggggggta	ggttaagtgc	cgggtaggca	aagtgggtga	cctcgatgct	11100
gcgtttcagc	acggctaggc	gcgcgttgct	accctcaagt	tcaccagca	ctccacagtg	11160
acttttcatt	tcgctgtttt	cttggttcag	agcgtttgag	gcgcgtttct	cgtcgcgtcc	11220
aagacctcca	aagatttttg	gcacttcgtc	gagcgaggcg	atatcaggta	tgacagcgcc	11280
ctgcgcgaag	gccagctgct	tgtccgctcg	gctgcggttg	gcacggcagg	ataggggtat	11340
cttgcaagtt	tggaaaaaga	tgtgataggt	ggcaagcacc	tctggcacgg	caaatacggg	11400
gtagaagttg	aggcgcgggt	tgggctcgca	tgtgccgttt	tcttggcggt	tgggggggtac	11460
gcgcggtgag	aacagggtgg	gttcgtaggc	aaggctgaca	tccgctatgg	cgagggggcac	11520
atcgctgcgc	tcttgcaacg	cgtcgcagat	aatggcgcac	tggcgctgca	gatgcttcaa	11580
cagcaagctg	tctccacat	ctaggtagtc	gccatgcctt	tgggtccccc	gcccgaactg	11640
ttcctcgctt	gcctctgctg	cgtcctgggt	ttgcttttta	tctctgtttg	gtactgagcg	11700
atcctcgctg	tcttcgctta	caaaacctgg	gtcctgctcg	ataatcactt	cctcctcctc	11760
aagcgggggt	gcctcgacgg	ggaagggtgg	aggcgcgttg	gcggcatcgg	tggaggcggt	11820
ggtggcggaac	tcaaaggggg	cggttaggct	gtcctccttc	tgcactgact	ccatgatctt	11880
tttctgccta	taggagaagg	aaatggccag	tccgggaagag	gagcagcgcg	aaaccacccc	11940
cgagcgcgga	cgcggtgcgg	cgcgacgtcc	accaaccatg	gaggacgtgt	cgcccccgct	12000
gccgtgcgcg	ccgcctcccc	gcgcgcccc	aaaaaagcgg	ctgaggcggg	gtctcgagtc	12060
cgaggacgaa	gaagactcgt	cacaagatgc	gctgggtgccc	cgcacaccca	gcccgcggcc	12120
atcgacctcg	acggcggatt	tggccattgc	gtocaaaaag	aaaaagaagc	gcccctctcc	12180
caagcccagag	cgcccggcat	ccccagaggt	gatcgtggac	agcgagggaag	aaagagaaga	12240
tgtggcgcta	caaatgggtg	gtttcagcaa	cccaccggtg	ctaataagc	acggcaaggg	12300
aggtaagcgc	acggtgcggc	ggctgaatga	agacgaccca	gtggcgcggg	gtatgcggac	12360
gcaagaggaa	aaggaagagt	ccagtgaagc	ggaaagtga	agcacgggtga	taaaccgct	12420
gagcctgcgc	atcgtgtctg	cgtgggagaa	gggcatggag	gctgcgcgcg	cgttgatgga	12480
caagtaccac	gtggataacg	atctaaaggc	aaacttcaag	ctactgcctg	accaagtgga	12540
agctctggcg	gccgtatgca	agacctgggt	aaacgaggag	caccgcgggt	tgcagctgac	12600
cttcaccagc	aacaagacct	ttgtgacgat	gatggggcga	ttcctgcagg	cgtacctgca	12660

gtcgttttgca	gaggtaacct	acaagcacca	cgagcccacg	ggctgcgcgt	tgtggctgca	12720
ccgctgcgct	gagatcgaag	gcgagcttaa	gtgtctacac	gggagcatta	tgataaataa	12780
ggagcacgctg	attgaaatgg	atgtgacgag	cgaaaacggg	cagcgcgcgc	tgaaggagca	12840
gtctagcaag	gccaagatcg	tgaagaaccg	gtggggccga	aatgtgggtg	agatctccaa	12900
caccgacgca	aggtgctgcg	tgcatgacgc	ggcctgtccg	gccaatcagt	tttccggcaa	12960
gtcttgccgc	atgttcttct	ctgaaggcgc	aaaggctcag	gtggctttta	agcagatcaa	13020
ggctttcatg	caggcgctgt	atcctaacgc	ccagaccggg	cacggtcacc	ttctgatgcc	13080
actacggtgc	gagtgcgaact	caaagcctgg	gcatgcaccc	tttttgggaa	ggcagctacc	13140
aaagttgact	ccgttcgccc	tgagcaacgc	ggaggacctg	gacgcggatc	tgatctccga	13200
caagagcgctg	ctggccagcg	tgacccaccc	ggcgctgata	gtgttccagt	gctgcaaccc	13260
tgtgtatcgc	aactcgcgcg	cgcaggggcg	aggccccaac	tgcgacttca	agatatcggc	13320
gcccgcacctg	ctaaacgcgt	tggatgatgg	gcgcagcctg	tggagtgaag	acttcaccga	13380
gctgccgcgg	atgggttgtg	ctgagtttaa	gtggagcact	aaacaccagt	atcgcaacgt	13440
gtccctgcca	gtggcgcata	gcgatgcgcg	gcagaacccc	tttgattttt	aaacggcgca	13500
gacggcaagg	gtggggggta	aataatcacc	cgagagtgtg	caaataaaaa	catttgcctt	13560
tattgaaagg	gtctcctagt	acattatttt	tacatgtttt	tcaagtgaac	aaaagaaagt	13620
gcgctcctaa	tctgcgcact	gtggctgcgc	aagtagggcg	agtggcgctc	caggaagctg	13680
tagagctgtt	cctgggttgc	acgcagggtg	ggctgtacct	ggggactgtt	aagcatggag	13740
ttgggtaccc	cggttaataag	gttcatggtg	gggttgtgat	ccatgggagt	ttggggccag	13800
ttggcaaagg	cgtggagaaa	catgcagcag	aatagtcac	aggcggccga	gttggggccc	13860
tgacgccttt	gggtggactt	ttccagcggt	atacagcggt	cgggggaaga	agcaatggcg	13920
ctacggcgca	ggagtgcact	gtactcaaac	tggtaaacct	gcttgagtcg	ttggtcagaa	13980
aagccaaagg	gctcaaagag	gtagcatgtt	tttgagcgcg	ggttccaggc	aaaggccatc	14040
cagtgtacgc	ccccagtcct	gcgaccggcc	gtattgacta	tggcgcaggc	gagcttgtgt	14100
ggagaaacaa	agcctggaaa	gcgcttgtca	taggtgccca	aaaaatatgg	cccacaacca	14160
agatctttga	caatggcttt	cagttcctgc	tcactggagc	ccatggcggc	agctgttgtt	14220
gatgttgctt	gcttctttta	tgttgtggcg	ttgccggccg	agaaggcgct	gcgcaggtag	14280
acggtctcga	tgacgcgcgc	gtgcggctgg	tgacacgcga	ccacgtcaaa	gacttcaaac	14340
aaaacataaa	gaagggtggg	ctcgctccatg	ggatccacct	caaaagtcac	gtctagcgcg	14400
tgggcggagt	tggcgtagag	aagggttttg	cccaggctct	tgagtgcgcc	catggacata	14460
aagttactgg	agaatgggat	gcgccaagg	gtgcgatcgc	aaagaaactt	tttctgggta	14520
atactgtcaa	ccgcggtttt	gcctattagt	gggtagggca	cgttggcggg	gtaagcctgt	14580
ccctcgcgca	tgggtgggagc	gaggtagcct	acgaatcctg	agttgttatg	ctggtgaaga	14640
attccaacct	cttgatactc	cttgatattt	gtatcgtcaa	ccacttgccg	gctcatgggc	14700
tggaaagttt	tgaagaacga	gtacatgcgc	tccttgtagc	tttctggaat	gtagaagccc	14760
tggtagccaa	tattgtagt	ggccaacatc	tgaccagga	accagtcctt	ggatcatgtg	14820
cactgagcta	cgttgtagcc	ctcccgcgtc	actgagcggt	taatctcaaa	ctcattggga	14880
gtaagcaggc	ggtcggttgc	cggccagcta	acagaagagt	caaaggtaat	ggccaccttc	14940
ttaaagggtg	gattaagata	gaagggtccg	tcaaggtagt	gtatggagcc	agagtagggtg	15000
tagtaagggt	ctagcctga	tcccagggaa	gggtttcctt	ttgtcttcaa	cgctgtgaag	15060
gccaaccgc	gaaatgctgc	ccagttgcgc	gatgggatgg	agatgggcac	gttgggtggcg	15120
ttggcggtga	tggggtatag	catgttggcg	gcggaaaggt	agtcattaaa	ggactgggtcg	15180
ttgggtgcat	ttctgagcat	ggcttccagc	gtggaggccg	tggttggggc	catgggggaag	15240
aagggtggcg	aaagacaaat	gctgtcaaac	ttaatgctag	ccccgtcaac	tctaagatcg	15300
tttcccagag	agctctgcag	aacctgttta	acatccttcc	tgaagttcca	ttcatatgta	15360
tatgagcctg	gcaggaggag	gaggttttta	atggcaaaaa	acttttgggg	cacctgaatg	15420
tgaaggggca	cgtagcggcc	gtttcccaac	aacatggagc	gataacggag	gccccgattg	15480
cggtgggtgg	taaagggtat	aacgttgtcc	atgtagtcca	gagaccagcg	cgccccaaag	15540
ttaatgtagc	agtctacaag	cccgggagcc	accactcgct	tgttcatgta	gtcgtagggtg	15600
ttgggggtgt	cagatatatt	cacattgggtg	gggttgtatt	ttagcttgtc	tggcagggtac	15660
agcgcaatat	tggagtaaag	gaaatttctc	cataggttgg	catttaggtt	aatttccatg	15720
gcaaagttgt	taccactctc	tatttcatta	cgtgttgcaa	aagtttcatc	ttttgtccat	15780
gtagtatctc	cattatcgcc	tgagccattg	ccattagcct	taatagcttg	ataggtgtca	15840
gttaccacca	tacccccaa	aggaaaacaa	taatttggca	attcatcctc	agttccatgg	15900
ttttcaatga	ttctaacatc	tggatcatag	ctgtctacag	cctgattcca	catagaaaaa	15960
tatctgggtt	tatcacctat	ggaatcaagc	aagagttgat	aggacagctc	tgtgtttctg	16020
tcttgcaaat	ctaccacggc	atttagctgc	gatgcctgac	cagcaagaac	accatgttg	16080
ccagtgcgtg	tataatacat	taggccaata	aaattgtccc	tgaaggaat	gtaattgggt	16140

ctgttttgga	tagattgttg	acccaacata	gcttttagaat	tttcatcacc	ttttccaggt	16200
ttgtaagaca	gatgtgtgtc	tgggggtttcc	atattttacat	cttccactgta	caaaaccact	16260
tttggttttag	tagcattgcc	ttgccgggtcg	ttcaaagagg	tagtattttga	gaagaattgc	16320
aagtcacact	ttggaagagg	cacccttttt	tcatccggaa	ccagaacgga	ttgaccacca	16380
aaaggatttg	taggcctggc	ataagatcca	tagcatgggt	tcatggggagt	tgtttttttta	16440
agcactctcc	ctcctgccgc	attagcatca	gcttcgtttcc	actgagattc	gccaatttga	16500
ggttcttggt	gataggaagg	atctgcgtat	acagggtttag	cttgtgtttc	tgcatgtgtc	16560
gacccatttt	gtagcccgct	ttttgttaatt	gtttctccag	acaaaggagc	ctgggcatag	16620
acatgtgttt	tcttagtagc	ctgatctcga	gcgtttttgtc	cttcttcttc	ctcttcttca	16680
tcttcatctt	cctcttcttc	atcctcggca	actgccggc	cgctatcttc	ggttttgttc	16740
cactcacagg	agttaggagc	gcccttggga	gctagagcgt	tgtaggcagt	gccggagtag	16800
ggcttaaaag	taggccccct	gtccagcacg	ccgcggtatgt	caaagtacgt	ggaagccata	16860
tcaagcacac	ggttgtcacc	cacagccagg	gtgaaccgcg	ctttgtacga	gtacgcggta	16920
tcctcgcggt	ccacagggat	gaaccgcagc	gtcaaacgct	gggaccggtc	tggtggttacg	16980
tcgtgcgtag	gtgccaccgt	ggggtttcta	aactttgttat	tcaggctgaa	gtacgtctcg	17040
gtggcgcggt	caaactgcac	cagcccggtg	ctcagggtact	ccgaggcgct	ctggcccgag	17100
atgtgcatgt	aagaccactg	cggcatcatc	gaaggggttag	ccatcttggga	aagcgggcgc	17160
acggcggttc	agcagctcct	ctggcggtga	catggacgca	tacatgacac	atacgacacg	17220
ttagctattt	agaagcatcg	tcggcgcttc	agggattgca	ccccagacc	cacgatgctg	17280
ttcagtggtg	tttgccagtt	gccactgggt	acgggcggca	tcgatcgcg	accgctggcg	17340
gcacggcgca	gggacgcgcg	gctagggcg	gttacaacaa	cggcgacg	ccctggcagc	17400
acaggtttct	gctgggtgtc	agcgggggga	gctaggtcca	gcgttacagg	tggtgtgtgg	17460
cccagcactc	cggtagccat	ggcgcggtg	ggaggggtgg	tggtcaggcc	ttgctttagt	17520
gcctcctcgt	acgagggagg	ctcatctatt	tgcttcacca	gagtttcttc	cctgtcgggc	17580
cgcgacgct	tttcgccacg	cccctctgga	gacactgtct	ccacggccgg	tggtggtctc	17640
tctacgggag	ggcggggatc	aagcttactg	ttaatcttat	tttgcactgc	ctggttggcc	17700
aggtccacca	ccccgcta	gccagaggcc	aggccatcta	ccaccttttg	ttggaaattt	17760
tgctctttca	acttgtccct	cagcatctgg	cctgtgtctg	tggtccaggc	cttgcgtcca	17820
tagttcttaa	tggtggaacc	gaaattttta	atgccgctcc	acagcgagcc	ccagctgaag	17880
gcgccaccgc	tcatattgct	ggtgccgata	tcttgccagt	ttcccatgaa	cggtcgcgag	17940
ccgtgtcgcg	gggccagaga	cgcaaagtgt	atgtcttcca	ttctacaaaa	tagttacagg	18000
accaagcgag	cgtgagactc	cagacttttt	attttgattt	ttccacatgc	aacttgtttt	18060
taatcagtg	ctctgcgcct	gcaaggccac	ggatgcaatt	ccgggcacgg	cgccaatcgc	18120
cgcgcgatc	agtggaaata	ggaggggcag	gataccgcgc	cgcatgcgac	ggtgcgacgc	18180
gcgcgcgcgc	cgggtggtgc	cacgacgc	gcgcgcgcgc	aggccgtggc	cggtcatgccc	18240
cctcctacgg	tgcatctctc	ctcggaatcc	cggtcaccgg	aaacggaggc	ggcaggtgag	18300
ggccatatct	gcaagaacca	caaagaccgg	cttttaaacc	atgctgggg	ggtagcgcg	18360
tggtggcagc	accagggctc	tgctcctctc	gcgagccacc	ctgcgcacgg	aaatcggggc	18420
cagcacgggc	tggtgcaggg	gacggcggtc	gcgggttcca	gtggtggttc	ggcgtcgggt	18480
agtcctctgt	cttctggggc	ggtaggtgta	gccacgtagc	ccgggggtag	gcgcgtgga	18540
aggatgtagg	gcataattcg	gcagtagtgc	gctggcggtg	ccgtacttcc	tggaacggcg	18600
cggtcgccgg	ggggctgaaa	cgcgaaacat	ccacgggtcc	gtttgcacct	ccgtagagg	18660
tttggaacgc	gccgcagcgg	ccgcctgcac	cgcgcatctc	gccaccgcgc	aggcaaccgc	18720
ggacgtttgt	gtctccatgc	cctctgtggc	agtggcaata	ctagtgtctac	tggtggtggg	18780
tatctgaacg	tccacggtct	gcacgcccac	tcccggtgct	acctgcttga	ttggccgcac	18840
gcggacctcg	ggctccagcc	caggctccac	ggtoattttt	tccaagacat	cttccagtcg	18900
ctggcgcttg	ggtaccatca	gctgcacggt	gggtgccaa	tcaccagact	cgcgcttttag	18960
gccgcgcttt	tcttcggacg	gtgcaagcgt	gggcagcacc	tgctgcagtg	tcacgggctt	19020
taggctagg	gttggtgtgc	cctcgtccag	cggcaaccgc	aacatgtcct	tatgccgctt	19080
tccgtaggca	aactccccga	ggcgctcgtt	ggcctgtctc	agcaggtcct	cgtcgccgta	19140
cacctcatca	tacacgcgct	tgtaggtgct	ggtggagcgc	tcaccgggcg	taaaaactac	19200
ggtggtgcgc	ggtcgcaaaa	cacgtcttac	gcgtgcacct	ttccactgta	ccgcgcgcct	19260
gggcgcgggt	gcgtgcagca	gttccacctc	gtcgtcaagt	tcatcatcat	catcatcttt	19320
ctttttcttt	ttgacctgct	ttagcttttc	gggcttgtaa	tcctgtctct	ccttcttcgg	19380
ggggccatag	atctccggcg	cgatgacctg	gagcatctct	tctttgattt	tgcgcttgga	19440
catagcttcg	ttgcgcgcgc	ccgcgcgtgg	atacatataa	cagtacgagt	ctaagtagtt	19500
ttttctttga	atctagttgc	gcggggggcg	ggtgcgcacg	ggcacgcgca	ggccgcta	19560
cgagtgcgcg	accaggtaca	cgttgccctc	gcgacctgca	gtcatagcac	taatggccgc	19620

ggctgctgcg	gcgcccgctc	gtcgccctgga	cctgggggggc	acagtgacaa	tacccgcggc	19680
cagccttcga	gcgcccgca	tggccgccc	tggccggtg	cgacgtgcg	ggttaagcag	19740
ggccgcccgc	gcgcgttggg	cggcagtgcc	gggtcgccgg	cggtggcgac	gtgctacgcg	19800
cctccgccgt	ctcttcattt	tagcataacg	ccgggtcccg	cgcaccacgg	tctgaatggc	19860
cgcgccact	gtggacactg	gtggcgccgt	gggcgtgtag	ttgcgcgcct	cctccaccac	19920
cgcgtcaatg	gcgtcatcga	cgggtggtgcg	cccagtgccg	ccgcgtttgt	gcgcgcccc	19980
gggcgcgcgg	tagtgcccgc	gcacgcgcac	tgggtgttg	tcggagcgct	tctttgcccc	20040
gccaacatc	ttgcttgga	agcgcaggcc	ccagcctgtg	ttattgctgg	gcgatataag	20100
gatggacatg	tttgctcaaa	aagtgcggct	cgataggacg	cgccggcaga	ctatgcccag	20160
ggccttgtaa	acgtaggggc	aggtgcggcg	tctggcgta	gtaatggtca	ctcgctggac	20220
tcctccgatg	ctgttgcgca	gcggtagcgt	cccgtgatct	gtgagagcag	gaacgttttc	20280
actgacggtg	gtgatggtgg	gggctggcgg	gcgcgccaaa	atctggttct	cgggaaagcg	20340
attgaacacg	tgggtcagag	aggtaaactg	gcggatgagc	tgggagtaga	cggcctggtc	20400
gttgtagaag	ctcttgaggt	gcacgggcaa	cagctcgccg	cccaccaccg	gaaagttgct	20460
gatctggctc	gtggagcggg	aggtcacggg	gtcttgcatc	atgtctggca	acgaccagta	20520
gacctgctcc	gagccgcagg	ttacgtcagg	agtgcaaaag	agggctccatg	agcggatccc	20580
gggtctgagg	tcgccgtagt	tgtatgcaag	gtaccagctg	cggtagctgg	tgaaggtgct	20640
gtcattgctt	attaggttgt	aactgcgttt	cttgctgtcc	tctgtcaggg	gtttgatcac	20700
cggtttcttc	tgaggcttct	cgacctcggg	ttgcgcagcg	ggggcgccag	cttctgccc	20760
tgccctcgcc	tcagcgcgct	tctcctccgc	ccgtgtggca	aaggtgtcgc	cgccaatggc	20820
atgatcgttc	atgtcctcca	ccggctgcct	tgcgcggcgt	gccgcgttgg	agttctcttc	20880
cgcgcgcgtg	ccactgttgt	tgcgcgcgcc	cgcgccatcc	ccgcctgtgt	cgggtgctac	20940
ttttaagctt	gcctggtagg	cgccacatc	caacagtgcg	ggaatgttac	cacctccag	21000
gtcatcgtag	gtgatcctaa	agccctcctg	gaagggttgc	cgcttgccga	tgcccaacaa	21060
gttgctcagg	cggctgtggg	tgaagtccac	cccgcctcct	ggcagcaaaa	tgatgtctgg	21120
atggaaggct	tcgtttgtat	ataccocagg	catgacaaga	ccagtgactg	ggtcaaaccc	21180
cagctctgaag	ttgcgggtgt	caaactttac	ccgatgtcgc	ctttccagaa	ccccgttctg	21240
cctgccacct	ttcaagtagt	gtccacgat	cgcgttgttc	ataaggtcta	tggtcatggt	21300
ctcggagtag	ttgccctcgg	gcagcgtgaa	ctccaccac	tcataatttca	gtccacctg	21360
tttgctccta	gtaagcgagc	gcgacaccat	caccgcgcgc	ttaaacttat	tggtaaacat	21420
gaactcgctc	acatttgcca	tggttggtatg	caggatgggt	ttcaggctgc	cgccccagtg	21480
cgaacggctg	tcaagattga	tggtctgtgt	gcttgccctc	cccgggctgt	agtcattgtt	21540
ttgaatgacc	gtggttagaa	agttgctgtg	gtcgttctgg	tagttcaggg	atgccacatc	21600
cgttgacttg	ttgtccacaa	ggtacacacg	ggtgggtgcg	aatagggttg	ccaactcaga	21660
gtaacggatg	ctgtttctcc	cccggtagg	ccgcaggtag	cgccggaggca	caaacggcgg	21720
gtccaggggg	gcacgaagg	gggaacccag	cgccgcgcgc	actggcgccg	cgctcaccac	21780
gctctcgtag	gagggaggag	gaccttcctc	atacatcgcc	gcgcgctgca	tactaagggg	21840
aatacaagaa	aaccaacgct	cgggtgccatg	gccttggtga	gttttttatt	ttgcatcatg	21900
ctttttttta	ttttttttta	aacattctcc	ccagcctggg	gcgaagggtg	gcaaacgggt	21960
tgccactccc	tcccaaatcc	aggacgtgc	tgtcgtctgc	cgagtcacgc	tcctccca	22020
ccagaccccg	ctgacggctc	tgcccttgac	gacgggtggg	cgggcgccgg	cggggcacat	22080
ccctgtgctc	ctgcgcatac	gtcttccatc	tactcatctt	gtccactagg	ctctctatcc	22140
cgttggtggg	aaatgccgga	ggcaggttct	tttcgcgctg	cggctgcagc	agcgagttgt	22200
ttagggtactc	ctcctcgccc	agcaggcgcg	ggcgggtggg	gcgagtgctg	gtaaaagacc	22260
ctatcaagct	tggaaatggg	ctactcgcat	ctgaccgcgg	ggccgcagcg	cctagatcgg	22320
acaagctgct	tggcctgcgg	aagctttcct	ttcgcagcgc	cgcctctgcc	tgctcgcgct	22380
gttgcaactc	tagcagggtc	tgccgttgcg	gggaaaacac	gctgtcgtct	atgtcgtccc	22440
agaggaatcc	atcgttaccc	tcgggcacct	caaatacccc	ggtgtagaaa	ccagggggcg	22500
gtagccagtg	cgggttcaag	atggcatttg	tgaaatactc	ggggttcacg	gcggccgcgc	22560
gatgcaagta	gtccattagg	cgattgataa	acggccgggt	tgaggcatac	atgcccgggt	22620
ccatggttcg	cgcggtcctg	tccagcgcca	cgtcggcgct	taccccgctg	cgcatcaggt	22680
taaggctcac	gctctgctgc	acatagcgca	agatgcgctc	ctcctcgctg	tttaaaactgt	22740
gcaacgaggg	gatcttctgc	cgccgggttg	tcagcaggta	gttcagggtt	gcctccaggg	22800
tgcccgctgc	ctcctgcccc	agcgcgcggc	tgacacttgt	aatctcctgg	aaagtatgct	22860
cgtccacatg	cgctgacct	atggcctcgc	ggtacagtg	cagcaagtga	cctaggtatg	22920
tgtcccggga	cacgctgcca	ctgtccgtga	agggcgctat	tagcagcagc	aacaggcgcg	22980
agttgggcgt	cagcaagcta	gacacggtcg	cgcgggtccc	tggtggagcc	cgcaccccc	23040
acagccctcg	caagtctctg	aaagcctggc	tcagggtttac	ggtctgcagg	ccttgtctac	23100

tgggtctggaa	aaaatagtct	ggccccggact	ggtacacctc	acttttgcggt	gtctcagtca	23160
ccatttagccg	cagtgcgctc	acaaagttgg	tgtagtccctc	ctgtccccgc	ggcacgttgg	23220
cgggctgtgt	actcaggaag	gcgttttagtg	caaccatgga	gccaggttgg	ccctgctgct	23280
gcgcgcgcctc	acgctgcgcc	acggcctcgc	gcacatcccc	caccagccgc	tccaggttgg	23340
tctgcacgtt	gccgctggtg	taacgagcca	cgcgctgaag	cagcgcgtcg	tagaccaggc	23400
cggcctcatc	gggcccgatg	gccctgtttt	cggccagcgc	gtttacgac	gccagcacct	23460
tctcgtgcgt	ggggtttgcg	cgcgcgcggga	ccaccgcttc	cagaattgcg	gagagccggt	23520
tggcctgcgg	ctgctgcggg	aacgcgtcag	ggttacgcgc	agtcagcgac	atgatgcggt	23580
ccatgacctg	gcgccagtcg	tccgtggagt	taaggccgga	cggctggctc	tgcagcgccg	23640
cccgcacccg	cgggtccggt	gcgtcttgca	tcattctgac	agaaacatca	ccgcttagta	23700
ctcgcgcgtcc	tctggctcgt	actcatcgctc	ctcgtcatat	tcctccacgc	cgccgacgtt	23760
gccagcgcgc	gcgggtgcca	ccgccagccc	aggctccggc	ccagctgcct	ccagggcgcg	23820
tccgcttggg	gccagcgca	ggtcagcgcc	cgcgtcaaag	taggactcgg	cctctctatc	23880
gccgctgccc	gtgccagcca	gggccctttg	caggtctgtgc	atcagctcgc	ggtcgttgag	23940
ctcgcgcgcg	cggctcacgc	tcacggcctt	gtggatgcgc	tcgttgcgat	aaacgcccag	24000
gtcgtcgctc	aaggtaagca	ccttcaacgc	catgcgcgtg	tagaacccct	cgatctttac	24060
ctccttgtct	atgggaacgt	aaggggtatg	gtatatcttg	cgggcgtaaa	acttgcccag	24120
actgagcatg	gaatagttaa	tggcggccac	cttgtcagcc	aggctcaagc	tgcgctcctg	24180
caccactatg	ctctgcagaa	tgtttatcaa	atcgagcagc	cagcggccct	cgggctctac	24240
tatgttttagc	agcgcacccc	tgaatgcctc	gttgtccctg	ctgtgctgca	ctataaggaa	24300
cagctgcgcc	atgagcggct	tgctattttg	gttttgcctc	agcgcgctta	caaagtccca	24360
cagatgcac	agtcctatag	ccacctcctc	gcgcgcacac	agcgtgcgca	cgtgggttgt	24420
aaagcttttt	tgaagtttaa	tctcctgggt	cacgcgtctgc	tcgtacgcgg	ttaccaggtc	24480
ggcggccgcc	acgtgtgcgc	gcgcgggact	aatcccggtc	cgcgcgtcgg	gctcaaagtc	24540
ctcctcgcgc	agcaaccgct	cgcggttcag	gccatgccgc	aactcgcgcc	ctgcgtggaa	24600
ctttcgatcc	cgcacctcct	cgggctcctc	tccctcgcgg	tcgcgaaaca	ggttctgcgg	24660
cggcacgtac	gcctcgcgcg	tgtcacgctt	cagctgcacc	cttgggtgtc	gctcaggaga	24720
gggcgcctct	agccgcgcca	ggccctcgcc	ctcctccaag	tccaggtagt	gccgggcccg	24780
gcgcgcgggg	ggttcgtaat	caccatctgc	cgcgcgcgtc	gccgcggatg	ttgcccctcc	24840
tgacgcggta	ggagaagggg	aggggtgccct	gcattgtctgc	cgctgctctt	gctcttgccg	24900
ctgctgagga	ggggggcgca	tctgccgcag	caccggatgc	atctgggaaa	agcaaaaaag	24960
gggctcgtcc	ctgtttccgg	aggaatttgc	aagcggggtc	ttgcatgacg	gggaggcaaa	25020
cccccgcttcg	ccgcagtcgg	gccggcccca	gactcgaaac	gggggtcctg	cgactcaacc	25080
cttggaaaat	aaccctccgg	ctacagggag	cgagccactt	aatgctttcg	ctttccagcc	25140
taaccgctta	cgcgcgcgcg	ggccagtggc	caaaaaagct	agcgcagcag	ccgcgcgcgc	25200
tggaaaggaa	ccaaaaggag	cgctcccccg	ttgtctgacg	tcgcacacct	gggttcgaca	25260
cgcgggcggg	aaccgcagtg	atcacggcgg	acggccggat	ccggggttcg	aaccgccggtc	25320
gtccgccatg	atacccttgc	gaatttatcc	accagaccac	ggaagagtgc	ccgcttacag	25380
gctctccttt	tgcacggctt	agagcgtcaa	cgactgcgca	cgctcacccg	gccagagcgt	25440
cccgactctt	gagcactttt	tgcgcgtcgc	caacatctgg	aaccgcgtcc	gcgactttcc	25500
gcgcgcctcc	accaccgcgg	ccggcatcac	ctggatgtcc	aggtacatct	acggatatca	25560
tcgccttatg	ttggaagacc	tcgcccccg	agccccggcc	accctacgct	ggccccctcta	25620
ccgcacggcg	ccgcgcgact	ttttgggtgg	atatcagtac	ctgggtgcgga	cttgcaacga	25680
ctacgtcttt	gactcaaggg	cttactcgcg	tctcagggtac	accgagctct	cgcagccggg	25740
tcaccagacc	gttaactggg	ccgttatggc	caactgcact	tacaccatca	acacgggcgc	25800
ataccacccg	tttgtggaca	tggatgactt	ccagctctacc	ctcacgcagg	tgcagcaggc	25860
catattagcc	gagcgcgttg	tcgcgcacct	ggccctgctt	cagccgatga	ggggcttcgg	25920
ggtcacacgc	atgggaggaa	gagggcgcca	cctacggcca	aactccgcgc	ccgcgcgtagc	25980
gatagatgca	agagatgcag	gacaagagga	aggagaagaa	gaagtgcggg	tagaaaggct	26040
catgcaagac	tactacaaag	acctgcgcgc	atgtcaaaaac	gaagcctggg	gcatggccga	26100
ccgcctgcgc	attcagcagg	ccggacccaa	ggacatggtg	cttctgtcga	ccatccgcgc	26160
tctcaagacc	gcctacttta	attacatcat	cagcagcacc	tccgccagaa	acaaccccga	26220
ccgccacccg	ctgccgcccc	ccacggtgct	cagcctacct	tcgcactgtg	actgggttaga	26280
cgcctttctc	gagaggtttt	ccgatccggg	cgatgcggac	tcgctcagg	ccctcgggtgg	26340
cggagtacct	acacaacaat	tgttgagatg	catcgtttagc	gccgtatccc	tgccgcacgg	26400
cagccccccg	ccaacccata	accgggacat	gacgggcggc	gtcttccaac	tgcccccccg	26460
cgagaacggc	cgcgcgcgtc	ccgagaccat	gcgcgcgtcg	cgcggggaga	tgatcgagcg	26520
ctttgtcgac	cgctcccg	tcgcgcgtcg	tcgcgcgcgt	gtccccctc	ccccaccgcc	26580

gccagaagaa	gaagaagaag	gggaggccct	tatggaagag	gagattgaag	aagaagaggc	26640
ccctgtagcc	tttgagcgcg	aggtgcgcga	cactgtcgcc	gagctcatcc	gtcttctgga	26700
ggaggagtta	accgtgtcgg	cgcgcaactc	ccagtttttc	aacttcgcgcg	tggacttcta	26760
cgaggccatg	gagcgccctg	aggccttggg	ggatatcaac	gaatccacgt	tgcgacgctg	26820
ggttatgtac	ttcttcgttg	cagaacacac	cgccaccacc	ctcaactacc	tctttcagcg	26880
cctgcgaaac	tacgccgtct	tcgcccggca	cgtggagctc	aatctcgcgc	aggtgggtcat	26940
gcgcgccccg	gatgccgaag	ggggcgtggt	ctacagccgc	gtctggaacg	agggaggcct	27000
caacgccttc	tcgcagctca	tggcccgcac	ctccaacgac	ctcgcgcgca	ccgtggagcg	27060
agccggagcg	ggagatctcc	aggaggaaga	gatcgagcag	ttcatggccg	aaatcgccta	27120
tcaagacaac	tcaggagacg	tgcaggagat	tttgcgccag	gcccgcgtca	acgacaccga	27180
aattgattct	gtcgaactct	ctttcagggt	caagctcacc	ggggcccgctg	tcttcacgca	27240
gaggcgccag	attcaggaga	tcaaccgcgcg	cgtcgtcgcg	ttcgccagca	acctccgcgc	27300
gcagcaccag	ctcctgcccc	cgcgcgggcg	cgacgtgccc	ctgccccctc	tcccggcggg	27360
tcccagagcc	cccttacctc	cgggggcccc	cccgcgtcac	cgtttttaga	tgcacatcc	27420
aaggacaccc	ccgcggcccc	ccgcccgcgcg	cgcggtaccg	tagtcgcgcg	gcggggatgc	27480
ggcctcttgc	agtcacatga	cgcgcgcacc	aaccagcccg	tggaaatcag	gtatcacctg	27540
gacctagccc	gcccctgac	cggcctatgc	gaggtaaacc	tgcagagct	cccgcctgac	27600
ctgtcgccgc	gggagctcca	gaccatggac	agctcccatc	tgcgcgatgt	tgtcatcaag	27660
ctccgacccg	cgcgcgcgga	catctggact	ttgggctcgc	gcggcgtggt	ggtccgatcc	27720
accataactc	ccctcgagca	gccagacggt	caaggacaag	cagccgaagt	agaagaccac	27780
cagccaaacc	cgccaggcga	ggggctcaaa	ttcccactct	gcttccttgt	gcgcggctcgt	27840
caggtcaaac	tcgtgcagga	tgtacagccc	gtgcacccgt	gccagtactg	cgcacgtttt	27900
tacaaaagcc	agcacgagtg	ttcggcccg	cgcagggact	tctactttca	ccacatcaac	27960
agccactcct	ccaactgggtg	gcgggagatc	cagttcttcc	cgatcggtc	gcacctcgc	28020
accgagcgtc	tctttgtcac	ctacgatgta	gagacctata	cttggatggg	ggcctttggg	28080
aagcagctcg	tgccttcat	gctggttatg	aagttcggcg	gagatgagcc	tctggtgacc	28140
gccgcgcgag	acctagccgt	ggaccttgga	tgggaccgct	gggaacaaga	cccgtcttacc	28200
ttctactgca	tcaccccaga	aaaaatggcc	ataggtcgcc	agtttaggac	ctttcgcgac	28260
cacctgcaaa	tgctaattggc	ccgtgacctg	tggagctcat	tcgtcgcttc	caaccctcat	28320
cttgcagact	gggccctgtc	agaacacggg	ctcagctccc	ctgaggagct	cacctacgag	28380
gaacttaaaa	aattgcccctc	catcaagggc	accccgcgct	tcttggaact	ttacatcgctg	28440
ggccacaaca	tcaacggctt	cgacgagatc	gtgctcgccg	cccaggtaat	taacaaccgt	28500
tccgaggtgc	cgggaccctt	ccgcatcaca	cgcaacttta	tgcctcgcgc	gggaaagata	28560
cttttcaacg	atgtcacctt	cgccctgcca	aaccgcggtt	ccaaaaagcg	cacggacttt	28620
ttgctctggg	agcagggcgg	atgcgacgac	actgacttca	aataccagta	cctcaaagtc	28680
atggttaggg	acacctttgc	gctcaccac	acctcgctcc	ggaaggccgc	gcaggcatac	28740
gcgtacccg	tagaaaagg	atgctgcgc	taccaggccg	tcaaccagtt	ctacatgcta	28800
ggctcttacc	gttcggaggc	cgacgggttt	ccgatccaag	agtactggaa	agaccgcgaa	28860
gagtttgttc	tcaaccgcga	gctgtggaaa	aaaaagggac	aggataagta	tgacatcatc	28920
aaggaaaccc	tggactactg	cgccctagac	gtgcaggcca	ccgcgcgact	cgtaacaag	28980
ctgcgcgact	cctacgcctc	cttcgtgcgt	gacgcggtag	gtctcacaga	cgccagcttc	29040
aacgtcttcc	agcgtccaac	catatcatcc	aaactcacatg	ccatcttcag	gcagatagtc	29100
ttccgagcag	agcagcccgc	ccgtagcaac	ctcgggtccc	acctcctcgc	tccctcgcac	29160
gaactatacg	attacgtgcg	cgccagcatc	cgcggtggaa	gatgctaccc	tacatatctt	29220
ggaatactca	gagagccctt	ctacgtttac	gacatttgcg	gcatgtacgc	ctccgcgctc	29280
acccacccca	tgccatgggg	tccccactc	aaccatacag	agcgcgcgct	tgccgcccgc	29340
gcatggcagc	aggcgctaga	cttgcaagga	tgcaagatag	actacttcga	cgcgcgccctg	29400
ctgccccggg	tctttaccgt	ggacgcagac	cccccggaag	agacgcagct	agacccacta	29460
ccgccattct	gttcgcgcaa	ggcgggccgc	ctctgctgga	ccaacgagcg	cctacgcgga	29520
gaggtagcca	ccagcgttga	ccttgtcacc	ctgcacaacc	gcggttggcg	cgtgcacctg	29580
gtgcccgcag	agcgcaccac	cgtctttccc	gaatggcggt	gcgttgccgc	cgaatacgtg	29640
cagctaaaca	tcgcggccaa	ggagcgcgcc	gatcgcgaca	aaaaccaaac	cctgcgctcc	29700
atcgccaagt	tgctgtccaa	cgccctctac	gggtcggttg	ccaccaagct	tgacaacaaa	29760
aagattgtct	tttctgacca	gatggacgcg	gccaccctca	aaggcatcac	cgcgggccag	29820
gtgaatatca	aatcctcctc	gtttttggaa	actgacaatc	ttagcgcaga	agtcatgccc	29880
gctttttgaga	gggagtactc	accccaacag	ctggccctcg	cagacagcga	tgcggaagag	29940
agtgaggacg	aacgcgcccc	cacccctttt	tatagccccc	cttcaggaac	acccggtcac	30000
gtggcctaca	cctataaaac	aatcaccttc	cttgatgcgcg	aagaggcgca	catgtgtctt	30060

cacaccctgg	agcgagtgg	ccccctagt	gacaacgacc	gctacccctc	ccacttagcc	30120
tccttcgtgc	tggcctggac	gcgagccttc	gtctcagagt	ggtccgagtt	tctatacgag	30180
gaggaccgcg	gaacaccgct	cgaggacagg	cctctcaagt	ctgtatacgg	ggacacggac	30240
agccttttcg	tcaccgagcg	tggacaccgg	ctcatggaaa	ccagaggtaa	gaaacgcata	30300
aaaaagcatg	ggggaaaacct	ggtttttgac	cccgaacggc	cagagctcac	ctggctcgtg	30360
gaatgcgaga	ccgtctgcgg	ggcctgcggc	gcggtatgct	actccccgga	atcgggtattt	30420
ctcgcgccca	agctctacgc	ccttaaaaagt	ctgcaactgc	cctcgtgcgg	cgctcctccc	30480
aagggcaagc	tgcgcgccaa	gggcccacgc	gcgagggggc	tggactatga	caccatgggtc	30540
aaatgctacc	tggccgacgc	gcagggcgaa	gaccggcagc	gcttcagcac	cagcaggacc	30600
agcctcaagc	gcacctgggc	cagcgcgag	cccggagcgc	accccttcac	cgtgacccag	30660
actacgctga	cgaggacctt	gcgcccgtgg	aaagacatga	ccctggcccg	tctggacgag	30720
caccgactac	tgcctgtacg	cgaaagccgc	cccaaccgcg	gaaacgagga	gatatgctgg	30780
atcgagatgc	cgtagagcac	gtgaccgagc	tgtgggaccg	cctggaactg	cttgggtcaaa	30840
cgctcaaaaag	catgcctacg	gcgagcggcc	tcaaaccgtt	gaaaaacttt	gcttccttgc	30900
aagaactgct	atcgctgggc	ggcgagcgcc	ttctggcgca	tttgggtcagg	gaaaacatgc	30960
aagtcaggga	catgcttaac	gaagtggccc	ccctgctcag	ggatgacggc	agctcagact	31020
ctcttaacta	ccagttgcag	ccggtaatag	gtgtgattta	cgggcccaac	cgctgcggta	31080
agtcgcagct	gctcaggaac	ctgctttctt	cccagctgat	ctccccctacc	ccggaacagg	31140
ttttcttcat	cgccccgcag	gtagacatga	tccccccatc	tgaactcaaa	gcgtgggaaa	31200
tgcaaatctg	tgagggtaac	tacgcccctg	ggccggatgg	aaccattata	ccgcagctctg	31260
gcaccctccg	cccgcgcttt	gtaaaaatgg	cctatgacga	tctcatcctg	gaacacaact	31320
atgacgttag	tgatcccaga	aatatcttcg	cccgccgcgc	cgcccggtgg	cccattggcca	31380
tcattatgga	cgaatgcatg	gaaaatctcg	gaggtcacaa	gggctgtctc	aagtctctcc	31440
acgcatttcc	ttctaagcta	catgacaaat	ttcccaagtg	caccggatac	actgtgctgg	31500
tggttctgca	caacatgaat	ccccggaggg	atatggctgg	gaacatagcc	aacctaaaaa	31560
tacagtccaa	gatgcctctc	atatccccac	gtatgcaccc	atcccagctt	aaccgctttg	31620
taaacactta	caccaagggc	ctgccccctg	caatcagctt	gctactgaaa	gacattttta	31680
ggcaccacgc	ccagcgctcc	tgctacgact	ggatcatcta	caacaccacc	ccgcagcatg	31740
aagctctgca	gtgggtgctac	ctccacccca	gagacgggct	tatgcccctg	tatctgaaca	31800
tccagagtca	cctttaccac	gtcctggaaa	aaatacacag	gaccctcaac	gaccgagacc	31860
gctgggtccc	ggcctaccgc	gcgcgcaaaa	cccctaaata	aagacagcaa	gacacttgct	31920
tgatcaaaa	ccaaacagag	tctgggtttt	atztatgttt	taaaccgcat	tgggagggga	31980
ggaagccttc	agggcagaaa	cctgctggcg	cagatccaac	agctgctgag	aaacgacatt	32040
aagttcccgg	gtcaaagaat	ccaattgtgc	caaaagagcc	gtcaacttgt	catcgcgggc	32100
ggatgaacgg	gaagctgcac	tgcttgcaag	cgggctcagg	aaagcaaagt	cagtcacaat	32160
cccgcgggcg	gtggctgcag	cggctgaagc	ggcgggcgag	gctgcagtct	ccaacggcgt	32220
tccagacacg	gtctcgtagg	tcaaggtagt	agagtttgcg	ggcaggacgg	ggcgaccatc	32280
aatgctggag	cccatcacat	tctgacgcac	cccggcccc	gggggcatgc	gcgttgtcaa	32340
atatgagctc	acaatgcttc	catcaaacga	tttggcgctc	atggcgggcg	ctgctgcaaa	32400
acagatacaa	aactacatga	gacccccacc	ttatatattc	tttcccacc	cttaagcccc	32460
cccatcgatg	gcaaacagct	attatgggta	ttatgggtgc	tagcgacatg	aggttgcccc	32520
gtattcagtg	tcgctgattt	gtattgtctg	aagttgtttt	tacgttaagt	tgatgcagat	32580
caattaatac	gatacctgcg	tcataattga	ttatttgacg	tggtttgatg	gcctccacgc	32640
acgttgtgat	atgtagatga	taatcattat	cacttttacgg	gtcctttccg	gtgatccgac	32700
aggttacggg	gcggcgacct	cgcgggtttt	cgctatttat	gaaaattttc	cggtttaagg	32760
cgtttcggtt	cttcttcgtc	ataacttaat	gtttttattt	aaaataacct	ctgaaaagaa	32820
aggaaacgac	agggtgctgaa	agcgaggctt	tttggcctct	gtcgtttcct	ttctctgttt	32880
ttgtccgtgg	aatgaacaat	ggaagttaac	ggatccaggc	cgcgagcaaa	aggccagcaa	32940
aaggccagga	accgtaaaaa	ggccgcgttg	ctggcgtttt	tccataggct	ccgccccctt	33000
gacgagcatc	acaaaaatca	acgctcaagt	cagaggtggc	gaaacccgac	aggactataa	33060
agataccagg	cgttttcccc	tggaaagctc	ctcgtgcgct	ctcctgttcc	gacctgccc	33120
cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	tggcgctttc	tcatagctca	33180
cgctgtaggt	atctcagttc	ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	33240
ccccccgttc	agcccagacc	ctgcgcctta	tccggtaact	atcgtcttga	gtccaaccgc	33300
gtaagacacg	acttatcgcc	actggcagca	gccactggta	acaggattag	cagagcgagg	33360
tatgtagggc	gtgctacaga	gttcttgaag	tgggtggccta	actacggcta	cactagaaga	33420
acagtatttg	gtatctgcgc	tctgccaaag	ccagttacct	tcggaaaaag	agttggtagc	33480
tcttgatccg	gcaaaacaaac	caccgctgggt	agcgggtggtt	tttttggttg	caagcagcag	33540

```

attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 33600
gctcagtggg acgaaaactc acgttaaggg attttggtca tcagattatc aaaaaggatc 33660
ttcacctaga tctttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag 33720
taaactttgt ctgacagtta ccaatgctta atcagtggag cacctatctc agcgatctgt 33780
ctatttcggt catccatagt tgccctgactc cccgtagtgt agataactac gatacgggag 33840
ggcttaccat ccggccccag tgctgcaatg ataccgcgtg acccacgctc accggctcct 33900
gatttatcag caataaacca gccagccgga agtgccgagc gcagaagtgg tcctgcaact 33960
ttatccgctt ccatccagtc tattagttgt tgccgggaag ctagagtaag tagttcgcca 34020
gttaatagtt ttcgcaacgt tggtgccatt gctacaggca tcgtggtgtc acgctcgtcg 34080
tttgggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc 34140
atgtttgtga aaaaagcggg tagctccttc ggtcctccga tagttgtcag aagtaagttg 34200
gccgcagtgt tatcactcat gggttatggca gcaactgcata attctcttac tgtcatgcca 34260
tccgtaagat gcttttctgt gactgggtgag tattcaacca agaatacggg ataataccgc 34320
gccacatagc agaactttaa aagtgtctcat cattgggaaa cgttcttcgg ggcgaaaact 34380
ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgcg caccgaagtg 34440
atcttctgca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa 34500
tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac ttttcctttt 34560
tcaatattat tgaagcattt atcagggtta ttgtctcctc agcggatata tattttg 34616

```

<210> 3

<211> 31672

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 3

```

gaatcggcca ggcggaatta actataacgg tcctaaggta gcgtcatcat cataatatac 60
cttatttttgg attgaagcca atatgataat gaggggggtg agtttgtgac gtggcgcggg 120
gcgtgggaac ggggcgggtg acgtaggttt tagggcggag taacttgcat gtattgggaa 180
ttgtagtttt tttaaaatgg gaagttacgt atcgtgggaa aacggaagtg aagatttgag 240
gaagtttgtg gtttttttggc tttcgtttct gggcgtaggt tcgcgtgcgg ttttctgggt 300
gttttttgtg gactttaacc gttacgtcat ttttttagtc tataataact cgctctgtac 360
ttggcccttt ttacactgtg actgattgag ctggtgccgt gtcgagtggg gttttttaat 420
aggttttttt actggttaagg ctgactgtta tggctgccgc tgtggaagcg ctgtatgttg 480
ttctggagcg ggagggtgct attttgccct ggcaggaggg tttttcaggt gtttatgtgt 540
ttttctctcc tattaatttt gttatacctc cttatggggc tgtaatgttg tctctacgcc 600
tgccgggtatg tattcccccg ggctatttct gtcgcttttt agcactgacc gatgttaacc 660
aacctgatgt gtttaccgag tcttacatta tgactccgga catgaccgag gaactgtcgg 720
tggtgctttt taatcacggg gaccagtttt tttacgggca cgcggcatg gccgtagtcc 780
gtcttatgct tataagggtt gtttttctct ttgtaagaca ggcttctaata gtttaaattg 840
ttttttttgt tatttttatt tgtgtttaat gcaggaaacc gcagacatgt ttgagagaaa 900
aatgggtgtc ttttctgttg tggttcggga acttaacctg ctttatctgc atgagcatga 960
ctacgatgtg cttgcttttt tgccgcaggc tttgcctgat tttttgagca gcaccttgca 1020
ttttatatcg ccgcccatgc aacaagctta cataggggct acgctgggta gcatagctcc 1080
gagtatgcgt gtcataatca gtgtgggttc ttttgtcatg gttcctggcg gggaagtggc 1140
cgcgctggtc cgtgcagacc tgcacgatta tgttcagctg gccctgcgaa gggacctacg 1200
ggatcgcggt atttttgtta atgttcgcgt ttggaatctt atacaggtct gtgagggaac 1260
tgaatttttg caatcatgat tcgctgcttg aggtgaaagg tggaggcgcg tctggagcag 1320
atttttacaa tggccggact taatatctgg gatttgctta gagacatatt gataaggtgg 1380
cgagatgaaa attatttggg catggttgaa ggtgctggaa tgtttataga ggagattcac 1440
cctgaagggt ttagccttta cgtccacttg gacgtgaggg cagtttgctt tttggaagcc 1500
attgtgcaac atcttacaaa tgccattatc tgttcttttg ctgtagagtt tgaccacgcc 1560
accggagggg agcgcgttca cttaatagat cttcattttg aggttttggg taatcttttg 1620
gaataaaaaa aaaaaaaaca tggttcttcc agctcttccc gtcctccccg tgtgtgactc 1680

```

gcagaacgaa	tgtgtagggt	ggctgggtgt	ggcttattct	gcggtggtgg	atgttatcag	1740
ggcagcggcg	catgaaggag	tttacataga	acccgaagcc	agggggcgcc	tggatgcttt	1800
gagagagtgg	atatactaca	actactacac	agagcgagct	aagcgacgag	accggagacg	1860
cagatctgtt	tgtcacgccc	gcacctgggt	ttgcttcagg	aaatatgact	acgtccggcg	1920
ttccatttgg	catgacacta	cgaccaacac	gatctcgggt	gtctcggcgc	actccgtaca	1980
gtagggatcg	cctacctcct	tttgagacag	agacccgcgc	taccatactg	gaggatcatc	2040
cgctgctgcc	cgaatgtaac	actttgacaa	tgcacaacgt	gagttacgtg	cgaggtcttc	2100
cctgcagtgt	gggattttacg	ctgattcagg	aatgggttgt	tccctgggat	atggttctga	2160
cgcgggagga	gcttgtaatc	ctgaggaagt	gtatgcacgt	gtgcctgtgt	tgtgccaaca	2220
ttgatatacat	gacgagcatg	atgatccatg	gttacgagtc	ctgggctctc	cactgtcatt	2280
gttccagttcc	cggttccctg	cagtgcatag	cggcggggca	ggttttggcc	agctggttta	2340
ggatgggtgg	ggatggcgcc	atgtttaatc	agaggtttat	atgggtaccg	gaggtgggtga	2400
attacaacat	gccaaaagag	gtaatgttta	tgtccagcgt	gtttatgagg	ggtcgccact	2460
taatctacct	gcgcttggtg	tatgatggcc	acgtgggttc	tgtggtcccc	gccatgagct	2520
ttggatacag	cgcttgccac	tgtgggattt	tgaacaatat	tgtggtgctg	tgctgcagtt	2580
actgtgctga	tttaagttag	atcagggtgc	ctactgtgct	ccggaggaca	agcgctctca	2640
tgctgcgggc	ggtgcgaatc	atcgctgagg	agaccactgc	catgttgtat	tcctgcagga	2700
cgagcggcg	gcggcagcag	tttattcgcg	cgctgctgca	gcaccaccgc	cctatcctga	2760
tgcacgatta	tgactctacc	cccatgtagg	cgtggacttc	cccttcgccc	cccgttgagc	2820
aaccgcaagt	tggacagcag	cctgtggctc	agcagctgga	cagcgacatg	aacttaagcg	2880
agctgccccg	ggagtttatt	aatatcactg	atcagcgttt	ggctcgacag	gaaaccgtgt	2940
ggaatataac	acctaagaat	atgtctgtta	cccatgatat	gatgcttttt	aaggccagcc	3000
ggggagaaag	gactgtgtac	tctgtgtgtt	gggagggagg	tggcagggtt	aatactaggg	3060
ttctgtgagt	ttgattaagg	tacggtgatc	aatataagct	atgtgggtgg	ggggctatac	3120
tactgaatga	aaaatgactt	gaaattttct	gcaattgaaa	aataaacacg	ttgaaacata	3180
acatgcaaca	ggttcacgat	tctttattcc	tgggcaatgt	aggagaaggt	gtaagagttg	3240
gtagcaaaa	tttcagtggt	gtatttttcca	ctttcccagg	accatgtaaa	agacatagag	3300
taagtgccta	ctcgctagt	ttctgtggat	tcactagtgc	cattaagtgt	aatggtaagt	3360
atcatagggt	tagttttatc	accatgcaag	taaacttgac	tgacaatgtt	attttttagca	3420
gtttgacttt	gggttttttg	ataggctaga	aggttaggca	taaatccaac	tgcattttgtg	3480
tatggatttg	cattagttga	gttcccattt	ctaaagttcc	agtaatgttt	tttaagttag	3540
gagttctcca	ttagaacacc	gttttggtca	aatctaagga	atatactaac	acttgcaacg	3600
gtgctgtgca	tggatgaaag	atctccagat	acagccaaag	cagctacagt	agctagtact	3660
tgactccac	attttgtaag	aaccaaagta	aatttgcagt	cattatctga	atgaattctg	3720
cagttaggag	atgggtcttg	ggttgtccac	agggtaagtt	tgtcatcatt	tttgtttcct	3780
attgtaatgg	cccctgagtt	gtcaaagctt	aaaccgcgtc	caagtttagt	aatcatggca	3840
ccgttttcat	tgtaatcaat	gccagagcca	attttagttt	ttattgggtt	gatatactgga	3900
gactcagatg	tgtttgtatc	aaactccaga	ccctttcctg	catttatagc	tatggcagta	3960
ttatcaaagt	ttagtccact	ggattttttt	atgctaactt	ccagtttttt	agtattgttt	4020
ttaatgcccc	aaaggatatag	gcctctgtta	tagtttatgt	ccaagttatg	atcctccatg	4080
atatacaggg	gtccctgccc	cagtttaaga	cgtagttttg	tttgagcatc	aaatgggtaa	4140
tcacacatcta	gaattaacaa	gttgttatatt	atacgcatgc	caccgcccgt	tttaattttcc	4200
atgttgtttg	atgaatcata	accaatagct	cctgcaactt	tggttctaag	ggagttttgt	4260
tcaacgggtga	cacctgggtcc	agtaactact	gttagtgtat	cggagttttg	tgctactttgc	4320
aaaggaccgc	ttattttaat	tcctattttt	ccattattta	cataaatagg	atcctccatg	4380
ttaatgcccc	agctacccgt	ggcagtagtt	agcgggggtg	atgcagttac	agtaagggtg	4440
tcgctgtcac	tgccagagag	gggggctgat	gtttgcaggg	ctagctttcc	atctgacact	4500
gtaatggggc	ctttagtagc	aatgcttagt	ttggagtctt	gcacggtcag	tggggcttgt	4560
gactgtacgc	taagagcgcc	gctagtaact	atcagaggag	cgggtggttc	cactgttagg	4620
gcgcctgagg	taattgtaag	tgggtgcggag	gtgtccaaac	ttatgtttga	ctttgttttt	4680
ttaagtggct	gagtaacagt	ggttacattt	tgggaggtga	ggtttccggc	cttgtctagg	4740
gtaagaccgc	tgcccatttt	aagcgcaagc	atgccgtggg	aggtgtccaa	aggttcggag	4800
acgcgtagag	agagaactcc	agggggactt	tcttggaaac	cattgggtga	aacaaatgga	4860
ggggtaagaa	agggcacagt	tggaggcccg	gtttctgtgt	catatggata	cacgggggtg	4920
aaggtgtctt	cagacggtct	ggcgcggttc	atctgcaaca	atatgaagat	agtggtgtcg	4980
gagggacaag	aacatgagga	atttgacatc	ccatttaaac	tttgagagaa	gtttgcagct	5040
aaaaggcggc	tgagatacca	gagttgggag	gaaggaaagg	aggtgatgct	gaataagctg	5100
gacaaagatt	tgctgactga	ttttaagtaa	gtaatttatt	cagtcgtagc	cgtccgccga	5160

gtctttcacc	gcgtcaaagt	tgggaataaa	ctgggtccggg	tagtggccgg	gaggtccaga	5220
aaaggggttg	aagtaaaccg	aaggcacgaa	ctcctcaata	aattgtagag	ttccaatgcc	5280
tccggagcgc	ggctccgagg	acgaggtctg	cagagttagg	atcgccctgac	ggggcgtaaa	5340
tgaagagcgg	ccagcgccgc	cgatctgaaa	tgtcccgtcc	ggacggagag	caagagagga	5400
gctcaccgac	tcgtcgttga	gctgaatacc	tcgcccctctg	attttcaggt	gagttatacc	5460
ctgccccggc	gaccgcaccc	tgtgacgaaa	gccgcccgcga	agctgcgccc	ctgagttagt	5520
catctgaact	tcggcctggg	cgctctctggg	aagtaccaca	gtggtggggag	cgggactttc	5580
ctggtacacc	agggcagcgg	gccaaactacg	gggattaagg	ttattacgag	gtgtggtggt	5640
aatagccgcc	tgttcgagga	gaattcgggt	tcgggtgggcg	cggattccgt	tgaccgcggga	5700
tatcatgttg	gggtcccgcgc	tcatgtagtt	tattcgggtt	gagtagtctt	gggcagctcc	5760
agccgcaagt	cccatttgtg	gctggttaact	ccacatgtag	ggcgtgggaa	tttcccttgc	5820
cataatggcg	ctgacgacag	gtgctggcgc	cgggtgtggc	cgctggagat	gacgtagttt	5880
tcgcgcttaa	atbtgagaaa	gggcgcgaaa	ctagtccctta	agagtcagcg	cgcagtattt	5940
gctgaagaga	gcctccgcgt	cttccagcgt	ggcgcgaagc	tgatctctgc	ttttgtgata	6000
caggcagctg	cgggtgaggg	agcgcagaga	cctgtttttt	attttcagct	cttgttcttg	6060
gcccctgctt	tggtgaaata	tagcatacag	agtggggaaa	atcctatttc	taagctcgcg	6120
ggtcgatacg	ggttcgttgg	gcgccagacg	cagcgctcct	cctcctgctg	ctgccgcgcg	6180
tgtggatttc	ttgggctttg	tcagagtctt	gctatccggg	cgcctttgct	tctgtgtgac	6240
cgctgctgtt	gctgcccgtg	ccgctgccgc	cgggtgcagta	ggggctgtag	agatgacggg	6300
agtaatgcag	gatgttacgg	gggaaggcca	cgccgtgatg	gtagagaaga	aagcggcggg	6360
cgaaggagat	gttgccccca	cagtcttgca	agcaagcaac	tatggcgctc	ttgtgcccgc	6420
gccacgagcg	gtagccttgg	cgctgttggg	gctcttgggc	taacggcgcg	ggctgcttag	6480
acttaccggc	cctggttcca	gtggtgtccc	atctacgggt	gggtcggcga	acaggcagtg	6540
ccggcggcgc	ctgaggagcg	gaggttgtag	cgatgctggg	aacggttgcc	aatttctggg	6600
gcgcgcggca	ggggaatgcg	accgaggggtg	acgggtgtttc	gtctgacacc	tcttcggcct	6660
cgggaagcttc	gtctaggctg	tcccagtcct	ccatcatctc	ctcctcctcg	tccaaaacct	6720
cctctgcctg	actgtcccag	tattcctcct	cgtccgtggg	tggcggcgcg	ggcagctgca	6780
gcttcttttt	gggtgccatc	ctgggaagca	agggcccgcg	gctgctgata	gggctgcggc	6840
ggcggggggga	ttgggttgag	ctcctcgccg	gactgggggt	ccaggtaaac	cccccgctcc	6900
tttcgtagca	gaaactcttg	gcgggctttg	ttgatggcct	gcaattggcc	aaggatgtgg	6960
ccctgggtaa	tgacgcaggc	ggtaagctcc	gcatttgccg	ggcgggattg	gtcttcgtag	7020
aacctaattct	cgtgggcgtg	gtagtcctca	ggtacaaatt	tgcgaaggta	agccgacgtc	7080
cacagccccg	gagtgagttt	caaccccggg	ggcgcggact	tttcgctcagg	cgagggaccc	7140
tgcagctcaa	aggtaccgat	aatttgactt	tgcttaagca	gttgcaattt	gcagaccagg	7200
gagcgggtgcg	gggtgcatag	gttgccagcga	cagtgacact	ccagtagggc	gtcaccgctc	7260
acgtcttcca	tgatgtcggg	gtggtaggca	aggtagtttg	ctagctgcag	aaggtagcag	7320
tgaccccaaa	gcggcgagg	gcattcacgg	tacttaattg	gcacaaagtc	gctaggaagc	7380
gcacagcagg	tggcgggcag	aattcctgaa	cgtcttagga	taaagtctct	aaagttttgc	7440
aacatgcttt	gactggtgaa	gtctggcaga	ccctgttgca	gggttttaag	caggcgatcg	7500
gggaagataa	tgtccgccag	gtgcgcggcc	cgttggaagg	cgttgaaggc	cgtccatagc	7560
tccttcaagt	tttgcttttag	cagcttctgc	agctccttta	ggttgcgctc	ctccaggcat	7620
tgctgccaca	cgcccatggc	cgtttgccag	gtgtagcaca	gaaataagta	aacgcagtcg	7680
cggacgtagt	cgcggcgccg	ctcgcccttg	agcgtggaat	gaagcacgtt	ttgcccaggg	7740
cggttttcgt	gcaaaattcc	aaggtagggg	accaggttgc	agagctccac	gttggaattt	7800
ttgcaggcct	ggcgcacgta	gccctggcga	aagggttagt	gcaacgtttc	ctctagcttg	7860
cgctgcatct	ccgggtcagc	aaagaaccgc	tgcatgcact	caagctccac	ggtaacaagc	7920
actgcggcca	tcattagctt	gcgtcgctcc	tccaagtcgg	caggctcgcg	cgtctcaagc	7980
cagcgcgcca	gctgctcatc	gccaaactgcg	ggtaggccct	cctcggtttg	ttcttgcaag	8040
tttgcatccc	tctccagggg	tcgtgcacgg	cgcacgatca	gctcgctcat	gactgtgctc	8100
ataaccttgg	ggggttaggt	aagtgcgggg	taggcaaaag	gggtgacctc	gatgctgcgt	8160
ttcagcacgg	ctaggcgcgc	gttgctaccc	tcaagttcca	ccagcactcc	acagtgactt	8220
tcattttcgc	tgttttcttg	ttgcagagcg	tttgccgcgc	gtttctcgtc	gcgtccaaga	8280
ccctcaaaga	tttttggcac	ttcgtcgagc	gaggcgatat	caggatatgac	agcgcctctg	8340
cgcaaggcca	gctgcttgct	cgctcggtcg	cggttggcac	ggcaggatag	gggtatcttg	8400
cagtttttga	aaaagatgtg	ataggtggca	agcactctctg	gcacggcaaa	tacggggtag	8460
aagttgaggg	gcgggttggg	ctcgcatgtg	cctgtttctt	ggcgtttggg	gggtacgcgc	8520
ggtgagaaca	gggtggcgctt	gtaggcaagg	ctgacatccg	ctatggcgag	gggcacatcg	8580
ctgcgcctct	gcaacgcgtc	gcagataatg	gcgcactggc	gctgcagatg	cttcaacagc	8640

acgtcgtctc	ccacatctag	gtagtcgcc	tgccttttgg	cccccgccc	gacttggtcc	8700
tgcgtttgc	ctgcgtcgtc	ctgggtcttg	tttttatcct	ctggttggtac	tgagcgatcc	8760
tgcgtcgtct	cgcttacaaa	acctgggtcc	tgcctcgataa	tcacttcctc	ctcctcaagc	8820
gggggtgcct	cgacggggaa	ggtggtaggc	gcgttgccgg	catcggtgga	ggcgggtggtg	8880
gcgaactcaa	agggggcggt	taggtgtgct	tccttctcga	ctgactccat	gatctttttc	8940
tgcctatagg	agaaggaaat	ggccagtcgg	gaagaggagc	agcgcgaaac	cacccccgag	9000
cgcggacgcg	gtgcggcgcg	acgtccacca	accatggagg	acgtgtcgtc	cccgtcgccg	9060
tgcgcgcgcg	ctccccgcgc	gcccccaaaa	aagcggctga	ggcggcgtct	cgagtcggag	9120
gacgaagaag	actcgtcaca	agatgcgctg	gtgccgcgca	cacccagccc	gcggccatcg	9180
acctcgacgg	cggattttggc	cattgcgtcc	aaaaagaaaa	agaagcggcc	ctctcccaag	9240
cccagcgcgc	cgccatcccc	agaggtgatc	gtggacagcg	aggaagaaag	agaagatgtg	9300
gcgctacaaa	tggtgggttt	cagcaaccca	ccggtgctaa	tcaagcacgg	caagggagggt	9360
aagcgcacgg	tgccggcggt	gaatgaagac	gacccagtg	cgccgggtat	gcggacgcaa	9420
gaggaaaagg	aagagtcag	tgaagcggaa	agtgaagaca	cggtgataaa	cccgtcgagc	9480
ctgccgatcg	tgtctgcgtg	ggagaagggc	atggaggctg	cgccgcgctt	gatggacaag	9540
taccacgtgg	ataacgatct	aaaggcaaac	tgcaagctac	tgccctgacca	agtggaaagt	9600
ctggcggcgc	tatgcaagac	ctggctaaac	gaggagcacc	gcgggttgca	gctgaccttc	9660
accagcaaca	agacctttgt	gacgatgatg	gggcgattcc	tgccagcgta	cctgcagtcg	9720
tttgcagagg	taacctacaa	gcaccacgag	cccacgggct	gcgcgttggtg	gctgcaccgc	9780
tgcgctgaga	tcgaaggcga	gcttaagtgt	ctacacggga	gcattatgat	aaataaggag	9840
cacgtgattg	aatggatgt	gacgagcgaa	aacgggcagc	gcgcgctgaa	ggagcagctc	9900
agcaaggcca	agatcgtgaa	gaaccgggtg	ggccgaaatg	tggtgcagat	ctccaacacc	9960
gacgcaagg	gctgcgtgca	tgacgcggcc	tgctcggcca	atcagttttc	cggaaggtct	10020
tgccgcatgt	tcttctctga	aggcgcaaa	gctcaggtgg	cttttaagca	gatcaaggct	10080
ttcatgcagg	cgctgtatcc	taacgcccag	accgggcacg	gtcaccttct	gatgccacta	10140
cggtgcgagt	gcaactcaaa	gcctgggcat	gcaccctttt	tggaaggcca	gctaccaaag	10200
ttgactccgt	tcgccctgag	caacgcggag	gacctggacg	cggatctgat	ctccgacaag	10260
agcgtgctgg	ccagcgtgca	ccaccggcgc	ctgatagtgt	tccagtgtcg	caacctgtgtg	10320
tatcgcaact	cgcgcgcgca	gggcgggagc	cccaactgcg	acttcaagat	atcggcgccc	10380
gacctgctaa	acgcgttgg	gatggtgcgc	agcctgtgga	gtgaaaactt	caccgagctg	10440
ccgcggatgg	ttgtgcctga	gtttaagtgg	agcactaaac	accagtatcg	caacgtgtcc	10500
ctgccagtg	cgcatagcga	tgccgcggcag	aacccctttg	atttttaaac	ggcgagagcg	10560
gcaagggtgg	ggggtaaata	atcaccgcag	agtgtacaaa	taaaaacatt	tgccctttat	10620
gaaagtgtct	ctagtacat	tattttttaca	tgtttttcaa	gtgacaaaaa	gaagtggcgc	10680
tcctaactctg	cgcactgtgg	ctgcggaaag	agggcgagtg	gcgctccagg	aagctgtaga	10740
gctgttcctg	gttgcgacgc	aggggtgggct	gtacctgggg	actgttaagc	atggagttgg	10800
gtaccccggt	aataagggttc	atgggtgggg	tgtgatccat	gggagtttgg	ggccagttgg	10860
caaaggcgtg	gagaaacatg	cagcagaata	gtccacaggc	ggccgagttg	ggccccctga	10920
cgctttgggt	ggactttttc	agcgtttatc	agcggtcggg	ggaagaagca	atggcgctac	10980
ggcgcaggag	tgactcgtac	tcaaactggt	aaacctgctt	gagtcggtgg	tcagaaaaagc	11040
caaagggtct	aaagaggtag	catgtttttg	agcgcgggtt	ccaggcaaa	gccatccagt	11100
gtacgcccc	agtctcgcga	ccggccgtat	tgactatggc	gcaggcgagc	ttgtgtggag	11160
aaacaaagcc	tggaagcgcc	ttgtcatagg	tgcccaaaaa	atatggccca	caaccaagat	11220
ctttgacaat	ggcttttcagt	tcctgtctac	tggagcccat	ggcggcagct	gttgtttgatg	11280
ttgcttgctt	ctttttatgtt	gtggcggttg	cggccgagaa	gggcgtgcgc	aggtacacgg	11340
tctcgatgac	gccgcgggtg	ggctggtgca	cacggaccac	gtcaaagact	tcaaacaaaa	11400
cataaagaag	ggtgggctcg	tccatgggat	ccacctcaaa	agtcatgtct	agcgcgtggg	11460
cggagtgggc	gtagagaagg	ttttggccca	ggtctgtgag	tgccgcccag	gacataaagt	11520
tactggagaa	tgggatgcgc	caaagggtgc	gatcgcaaa	aaactttttc	tgggtaatac	11580
tgtcaaccgc	ggtttttgct	attagtgggt	agggcacgtt	ggcggggtaa	gcctgtccct	11640
cgcgcgatgt	gggagcgagg	tagcctacga	atcctgagtt	gttatgctgg	tgaagaattc	11700
caacctgctg	atactccttg	tatttagtat	cgtcaaccac	ttgccggctc	atgggctgga	11760
agtttctgaa	gaacgagtac	atgcggctct	tgtagctttc	tggaatgtag	aagccctggt	11820
agccaatatt	gtagttggcc	aacatctgca	ccaggaacca	gtccttggtc	atgttgcaact	11880
gagctacgtt	gtagccctcc	ccgtcaactg	agcgtttaat	ctcaaactca	ttgggagtaa	11940
gcaggcgggt	ggtgcccggc	cagctaacag	aagagtcaaa	ggtaatggcc	accttcttaa	12000
aggtgtgatt	aagatagaag	gttccgtcaa	ggtatggtat	ggagccagag	taggtgtagt	12060
aagggtcgta	gcctgatccc	aggggaaggg	tttcccttgg	cttcaagcgt	gtgaaggccc	12120

aaccgcgaaa	tgctgcccag	ttgcgcgatg	ggatggagat	gggcacgttg	gtggcggttg	12180
cgggtatggg	gtatagcatg	ttggcggcgg	aaaggtagtc	attaaaggac	tggtcggttg	12240
tgtcatttct	gagcatggct	tccagcgtgg	aggccgtggt	gtgggccatg	gggaagaagg	12300
tggcgtaaag	acaaatgctg	tcaaacttaa	tgctagcccc	gtcaactcta	agatcgtttc	12360
ccagagagct	ctgcagaacc	atgttaacat	ccttccctgaa	gttccattca	tatgtatatg	12420
agcctggcag	gaggaggagg	tttttaattg	caaaaaactt	ttggggcacc	tgaatgtgaa	12480
agggcacgta	gcggccggtt	cccaacaaca	tggagcgata	acggaggccc	gcattgcggt	12540
ggtggttaaa	gggattaacg	ttgtccatgt	agtccagaga	ccagcgcgcc	ccaagggttaa	12600
tgtagcagtc	tacaagcccg	ggagccacca	ctcgcttggt	catgtagtcg	taggtgtttg	12660
ggttgtcaga	tatttccaca	ttggtggggg	tgtatttttag	cctgtctggc	aggtacagcg	12720
caatatttga	gtaaaggaaa	tttctccata	ggttggcatt	taggttaatt	tccatggcaa	12780
agttgttacc	cactcctatt	tcattacgtg	ttgcaaaagt	ttcatctttt	gtccatgtag	12840
tatctccatt	atcgccctgag	ccattgccat	tagccttaat	agccttgatag	gtgtcagtta	12900
ccccaatacc	ccaagagga	aaacaataat	ttggcaattc	atcctcagtt	ccatggtttt	12960
caatgattct	aacatctgga	tcatagctgt	ctacagcctg	attccacata	gaaaaatatc	13020
tggttctatc	acctatggaa	tcaagcaaga	ggtgatagga	cagctctgtg	tttctgtctt	13080
gcaaactctac	cacggcattt	agctgcgatg	cctgaccagc	aagaacaccc	atgttgccag	13140
tgctgttata	at'acattagg	ccaataaaat	tgtccctgaa	agcaatgtaa	ttgggtctgt	13200
ttggcataga	ttgttgaccc	aacatagctt	tagaattttc	atcacctttt	ccaggtttgt	13260
aagacagatg	tgtgtctggg	gtttccatat	ttacatcttc	actgtacaaa	accacttttg	13320
gtttagtagc	attgccttgc	cggctcgttca	aagaggtagt	atttgagaag	aattgcaagt	13380
caaccttttg	aagaggcacc	cctttttcat	ccggaaccag	aacggattga	ccaccaaaag	13440
gattttgtagg	cctggcataa	gatccatagc	atggtttcat	gggagttgtt	tttttaagca	13500
ctctccctcc	tgcgcatta	gcatacagct	cgttccactg	agattcgcca	atttgagggt	13560
ctggttgata	ggaaggatct	gcgtatacag	gttttagcttg	tgtttctgca	ttgtctgata	13620
ctattttgtag	cccgcctttt	gtaattgttt	ctccagacaa	aggagcctgg	gcatagacat	13680
gtgttttctt	agtagcctga	tctcgagcgt	tttgtctctt	ttcttctctt	tttctctctt	13740
catcttctct	ttcttcatcc	tccgcaactg	cccgcccgct	atcttcggtt	tgttccctact	13800
cacaggagtt	aggagcgccc	ttgggagcta	gagcgttgta	ggcagtgccg	gagtagggct	13860
taaaagtagg	ccccctgtcc	agcacgccgc	ggatgtcaaa	gtacgtggaa	gccatatcaa	13920
gcacacgggt	gtcaccacaca	gccagggtga	accgcgcttt	gtacgagtac	gcggtatcct	13980
cgcggtccac	agggatgaac	cgcagcgtca	aacgctggga	ccggtctgtg	gttacgtcgt	14040
gcgtagggtc	caccgtgggg	tttctaaact	tgttattcag	gctgaagtac	gtctcggttg	14100
cgcgggcaaa	ctgcaccagc	ccgggggtca	ggtactccga	ggcgtcctgg	cccagatgtg	14160
gcattgtaaga	ccactgcggc	atcatcgaag	gggtagccat	cctggaaagc	gggcgcacgg	14220
cggctcagca	gctcctcttg	cggcgacatg	gacgcataca	tgacacatac	gacacgttag	14280
ctatttagaa	gcatacgtcg	cgcttcaggg	attgcacccc	cagacccacg	atgctgttca	14340
gtgtgctttg	ccagttgcca	ctggctacgg	gccgcacgca	tgcgggaccg	ctggcgccac	14400
ggcgcaggga	cgcgcggcta	gggcgggtta	caacaacggc	ggacggccct	ggcagcacag	14460
ggttctctct	ggtgtcagcg	gggggaggca	ggctccagct	tacaggtgtg	tgctggccca	14520
gcaactccggt	agccatgggc	gcgatgggac	gggtggtggg	caggccttgc	tttagtgctt	14580
cctcgtacga	gggaggctca	tctattttgc	tcaccagagt	ttcttccctg	tcgggcgcgc	14640
gacgcttttc	gccacgcccc	tctggagaca	ctgtctccac	ggcgggtgga	ggctcctcta	14700
cgggaggggc	gggatcaagc	ttactgttaa	tcttattttg	cactgcctgg	ttggccagggt	14760
ccaccacccc	gctaattgca	gaggccaggc	catctaccac	cctttgtttg	aaatttttgt	14820
ccttcaactt	gtccctcagc	atctggcctg	tgtgtctgtt	ccaggccttg	ctgccatagt	14880
tcttaattggt	ggaaccgaaa	tttttaattg	cgtctccacag	cgagccccag	ctgaaggcgc	14940
cacgcgtcat	attgctggtg	ccgatatact	gccagtttcc	catgaacggg	cgcgagccgt	15000
gtcgcggggc	cagagacgca	aagttgatgt	cctccattct	acaaaatagt	tacaggacca	15060
agcgagcgtg	agactccaga	cttttttatt	tgattttttc	acatgcaact	tgttttttaa	15120
cagtgtctct	gcgcctgcaa	ggccacggat	gcaattccgg	gcacggcgcc	aatcgccgcg	15180
gcgatcagtg	gaataaggag	gggcaggata	ccgcgcgcga	tgcgacggtg	cgacgcgcgc	15240
cgcgcgcggt	ggtgcgcacg	acgcattgcc	cccgtcaggc	cgtggccggc	catgcccctc	15300
ctacggtgca	ttcttccctc	gaatcccggc	accgggaaac	ggaggcggca	ggtgagggcc	15360
atatctgcaa	gaaccacaaa	gaccggcttt	taaacgatgc	tggggtggta	gcgcgctgtt	15420
ggcagcacca	gggtcctgcc	tccttcgcga	gccaccctgc	gcacggaaat	cgggggccagc	15480
acgggctggc	gacggcgacg	gcggcgggcg	gttcagtggt	tggttcggcg	tcgggtagtc	15540
gctcgtcttc	tggggcggta	ggtgtagcca	cgatagccgg	gggtaggcgc	gatggaagga	15600

tgtagggcat	attcgggag	tagtgcgctg	ggggtgcgct	acttcctgga	acggcgcgggg	15660
cgccgggggg	ctgaaacgcg	aaacatccac	gggtccggtt	gcacctccgt	agaggttttg	15720
gacgcgccg	cagcgccg	ctgcaccg	gcatctgcca	ccgccgaggc	aaccggggac	15780
gtttgtgtct	ccatgccctc	tgtggcagtg	gcaatactag	tgctactggg	ggtgggtatc	15840
tgaacgtcca	cggtctgcac	gcccagtc	ggtgccacct	gcttgattgg	ccgcacgcgg	15900
acctcgggct	ccagcccagg	ctccacgggt	atTTTTTcca	agacatcttc	cagtcgctgg	15960
cgcttggtta	ccatcagctg	cacgggtgggt	gccaagtcac	cagactcgcg	ctttaggccg	16020
cgcttttctt	cggacgggtg	aagcgtgggc	agcacctgct	gcagtgtcac	gggcttttag	16080
ctaggtgttg	ggttgccctc	gtccagcggc	aacgccaaaca	tgctccttat	ccgctttccg	16140
tagggcaact	ccccgaggcg	ctcgttggcc	tgctcaagca	ggtcctcgct	gccgtacacc	16200
tcatacatac	cgcgcttgta	ggtgcgggtg	gagcgctcac	cgggcgtaaa	aactacgggtg	16260
gtgcggggtc	gcaaaacacg	tcttacgcgt	cgacctttcc	actgtaccgg	ccgcctgggc	16320
gcggttgcg	gcagcagttc	cacctcgctg	tcaagttcat	catcatcatc	atctttcttt	16380
ttctttttga	cccgcttttag	ctttcggggc	ttgtaatcct	gctcttcctt	cttcgggggg	16440
ccatagatct	ccggcgcgat	gacctggagc	atctcttctt	tgattttg	cttggaata	16500
gcttctgttc	gcgcgcgcgc	cgctggatac	atacaacagt	acgagtctaa	gtagtttttt	16560
cttgcaatct	agttgcgcgg	ggggcggggtg	cgcacgggca	cgcgaggcc	gctaaccgag	16620
tcgcgacccc	agtacacgtt	gcccctgcga	ccctgagtc	tagcactaat	ggcgcgggct	16680
gctgcggcg	ccgctcgctg	cctggacctg	gggggcacag	tgacaatacc	cgcgcccgag	16740
cttcgagcgg	cccgcatggc	cgcccgctcg	ccggtgcgac	gtgcgcgggt	aagcagggcc	16800
gcccgcgcgc	gttgggcggc	agtgcgggt	cgggcgcggt	ggcgacgtgc	tacgcgcctc	16860
cgccgtctct	tcatttttag	ataacgcgg	gctccgcgca	ccacggctctg	aatggcccg	16920
tcactgttgc	acactgggtg	cgcgctgggg	gtgtgattgc	gcgcctcctc	caccaccgcg	16980
tcaatggcgt	catcgacgg	ggtgcgcaca	gtgcggcg	gtttgtgcgc	gcccaggggc	17040
gcgcggtagt	gcccgcgcac	gcgcactggg	tgttggtcgg	agcgcttctt	tgccccgcca	17100
aacatcttgc	ttgggaagcg	caggccccag	cctgtgttat	tgctgggcca	tataaggatg	17160
gacatgtttg	ctcaaaaagt	gcggctcgat	aggacgcgcg	gcgagactat	gcccagggcc	17220
ttgtaaacgt	aggggcaggt	gcggcgctcg	gcgtcagtaa	tggtcactcg	ctggactcct	17280
ccgatgctgt	tgcgacgcgg	tagcgtcccg	tgatctgtga	gagcaggaac	gttttactctg	17340
acgggtggtga	tggtgggggc	tgggcgggcg	gcaaaaatct	ggttctcggg	aaagcgattg	17400
aacacgtggg	tcagagaggt	aaactggcg	atgagctggg	agtagacggc	ctggctcgttg	17460
tagaagctct	tggagtgcac	gggcaacagc	tcggcgccca	ccaccggaaa	gttgctgata	17520
tggtctcggtg	agcgggaagg	cacggggtct	tgcatcatgt	ctggcaacga	ccagtagacc	17580
tgctccgagc	cgcaggttac	gtcaggagtg	caaaggagg	tccatgagcg	gatccccgtc	17640
tgagggtcgc	cgtagttgta	tgcaaggtag	cagctgcgg	actgggtgaa	ggtgtctca	17700
ttgcttatta	ggttgtaact	gcgtttcttg	ctgtcctctg	tcagggggtt	gatcacgggt	17760
ttcttctgag	gcttctcgac	ctcgggttgc	gcagcggggg	cggcagcttc	tgccgctgcc	17820
tcggcctcag	cgcgcttctc	ctccgcccgt	gtggcaaaag	tgctcgccgcg	aatggcatga	17880
tcgttcatgt	cctccaccgg	ctgcattgcc	cgcgctgccg	cgttggagtt	ctcttcgcgcg	17940
ccgctgccac	tggtgttgcc	gccgcctgcg	ccatccccgc	cctgttcggg	gtcacttttt	18000
aagcttgcc	gtaggcgtc	cacatccaac	agtgcgggaa	tgttaccacc	ctccagggtca	18060
tcgtagggtga	tcctaaagcc	ctcctggaag	ggttgccgct	tgcggtatgc	caacaagttg	18120
ctcaggcggc	tggtgggtgaa	gtccaccccg	catcctggca	gcaaaatgat	gtctggatgg	18180
aaggcttcgt	ttgtatatac	cccaggcatg	acaagaccag	tgactgggtc	aaaccccagt	18240
ctgaagttgc	gggtgtcaaa	ctttaccccg	atgtcgcttt	ccagaacccc	gttctgcctg	18300
cccacttttc	agtagtgctc	cacgatcgcg	ttgttcataa	ggtctatggt	catggtctcg	18360
gagtagttgc	cctcgggcag	cgtgaactcc	acccactcat	atctcagctc	cacctgtttg	18420
tccttagtaa	gcgagcgcg	caccatcacc	cgcgcttaa	acttattggt	aaacatgaac	18480
tcgttcacat	ttggcatgtt	ggtatgcagg	atggttttca	ggtcgccgc	ccagtgcgaa	18540
cggtcgtaaa	gattgatggt	ctgtgtgctt	gcctcccccg	ggctgtagtc	attgttttga	18600
atgaccgtgg	ttagaaagtt	gctgtggtcg	ttctggtagt	tcagggatgc	cacatccgtt	18660
gacttggtgt	ccacaaggta	cacacgggtg	gtgtcgaata	gggtgcca	ctcagagtaa	18720
cggtatgctgt	ttctcccccc	ggtaggccgc	aggtaccgcg	gaggcaciaa	cgggcggtcc	18780
aggggagcat	cgaaggggga	acccagcgcc	gcgcgcactg	gcgcgcgcgt	caccacgctc	18840
tcgtaggagg	gaggaggacc	ttcctcatac	atcgccgcgc	gctgcatact	aaggggaata	18900
caagaaaacc	aacgctcggt	gccatggcct	tggtgagttt	tttattttgc	atcatgcttt	18960
tttttttttt	ttttaaaaca	ttctccccag	cctggggcga	aggtgcgcaa	acgggttgcc	19020
actccctccc	aatccaggga	cgctgctgtc	gtctgcccag	tcctcgtcct	cccacaccag	19080

accccgctga	cggtcgtgcc	tttgacgaag	ggtggggcggg	cgcggggcggg	gcacatccct	19140
gtgctcctgc	gcatacgtct	tccatctact	catcttggtcc	actaggctct	ctatcccgtt	19200
gttgggaaat	gccggaggca	ggttcttttc	gcgctgcggc	tgcagcagcg	agttgttttag	19260
gtactcctcc	tcgcccagca	ggcgcgggcg	ggtggtgcga	gtgctggtaa	aagaccctat	19320
caagcttgga	aatgggctac	tcgcatctga	ccgcggggcg	gcagccgcta	gacgggacaa	19380
gctgcttggc	ctgcggaagc	tttcctttcg	cagcgccggc	tctgcctgct	cgcgctgttg	19440
caactctagc	agggtctgcg	ggtgcgggga	aaacacgctg	tcgtctatgt	cgtcccagag	19500
gaatccatcg	ttaccctcgg	gcacctcaaa	tcccccggtg	tagaaaccag	ggggcggtag	19560
ccagtgcggg	ttcaagatgg	cattggtgaa	atactcgggg	ttcacggcgg	ccgcgcgatg	19620
caagtagtcc	attaggcgat	tgataaacgg	ccggttttag	gcatacatgc	ccggttccat	19680
gttgcgcgcg	gtcatgtcca	gcgccacgct	gggcgttacc	ccgtcgcgca	tcaggttaag	19740
gctcacgctc	tgctgcacat	agcgcaagat	gcgctcctcc	tcgctgttta	aactgtgcaa	19800
cgaggggatc	ttctgcgcgc	ggttggtcag	caggtagttc	agggttgctc	ccaggctgcc	19860
cgtgtcctcc	tgccccagcg	cgcggtcgac	acttgtaatc	tcctggaaag	tatgctcgtc	19920
cacatgcgcc	tgacctatgg	cctcgcggta	cagtgtcagc	aagtgcacta	ggtatgtgtc	19980
ccggggacacg	ctgccactgt	ccgtgaaggg	cgctattagc	agcagcaaca	ggcgcgagtt	20040
ggcggtcagc	aagctagaca	cggtcgcgcg	ggtccctgtg	ggagcccgca	ccccccacag	20100
cccttgcaag	ttcttgaaag	cctgggtcag	gtttacggtc	tgcaggcctt	gtctactggt	20160
ctggaaaaaa	tagtctggcc	cggactggta	cacctcactt	tgcggtgtct	cagtcaccat	20220
tagccgcagt	gcgctcacaa	agttggtgta	gtcctcctgt	ccccgcggca	cgttggcggg	20280
ctgtgtactc	aggaaggcgt	ttagtgcac	catggagccc	aggttgccct	gctgctgcgc	20340
gcgctcacgc	tgcgccacgg	cctcgcgcac	atccccacc	agccggtcca	ggttggtctg	20400
cacgttgccg	ctggtgtaac	gagccacgcg	ctgaagcagc	cgctcgtaga	ccaggccggc	20460
ctcatcgggc	cggatggccc	tgttttcggc	cagcgcgttt	acgatcgcca	gcaccttctc	20520
gtgcgtgggg	tttgcgcgcg	ccgggaccac	cgcttcacga	attgcggaga	gcgggttggc	20580
ctgcggctgc	tgccggaacg	cgtcagggtt	acgcgcagtc	agcgacatga	tgcggtccat	20640
gacctggcgc	cagtcgtccg	tggagttaag	gccggacggc	tggctctgca	gcgcgcgccg	20700
caccgcgggg	tccgttgctg	cttgcatcat	ctgatcagaa	acatcacgcg	ttagtactcg	20760
ccgtcctctg	gctcgtaact	atcgtcctcg	tcataattcct	ccacgcgcgc	ccaggttgcca	20820
gcgcgcgcgc	gtgccaccgc	cagcccagggt	ccggccccag	ctgcctccag	ggcgcgctcg	20880
cttggggccc	agcgcgagtc	agcgcccgcg	tcaaagtagg	actcggcctc	tctatcgccg	20940
ctgcccgctg	cagccagggc	cctttgcagg	ctgtgcatca	gctcgcggtc	gctgagctcg	21000
cgccgcgggc	tcacgctcac	ggccttggtg	atgcgctcgt	tgcgataaac	gcccaggctcg	21060
tcgctcaagg	taagcacctt	caacgccatg	cgcagttaga	acccctcgat	ctttacctcc	21120
ttgtctatgg	gaacgtaagg	ggtatggtat	atcttgccgg	cgtaaaactt	gcccagactg	21180
agcatggaat	agttaatggc	ggccaccttg	tcagccaggc	tcaagctgcg	ctcctgcacc	21240
actatgctct	gcagaatgtt	tatcaaatcg	agcagccagc	ggccctcggg	ctctactatg	21300
tttagcagcg	catccctgaa	tgccctggtg	tccctgctgt	gctgcactat	aaggaacagc	21360
tgcgccatga	gcggcttgct	atgttggttt	tgctccagcg	cgcttacaaa	gtcccacaga	21420
tgcatacgtc	ctatagccac	ctcctcgcgc	gccacaagcg	tgcgcacgtg	ggtgttaaag	21480
cttttttgaa	agttaatctc	ctggttccac	gtctgctcgt	acgcggttac	caggtcgggc	21540
gccgccacgt	gtgcgcgcgc	gggaactaat	ccggtccgcg	cgctcgggctc	aaagtcctcc	21600
tcgcgcagca	accgctcgcg	gttcaggcca	tgccgcaact	cgcgccttgc	gtggaacttt	21660
cgatcccgcg	tctcctcggg	ctcctctccc	tcgcggtcgc	gaaacagggt	ctgccgcggc	21720
acgtacgcct	cgcgcgtgtc	acgcttcagc	tgcaccttg	ggtgtcgtc	aggagagggc	21780
gctcctagcc	gcgccaggcc	ctcgccctcc	tccaagtcca	ggtagtgcgc	ggcccggcgc	21840
cgcggggggt	cgtaatcacc	atctgcgcgc	gcgtcagccg	cggatggttg	ccctcctgac	21900
gcggtaggag	aaggggaggg	tgccctgcat	gtctgcccgt	gctcttgctc	ttgcgcgtgc	21960
tgaggagggg	ggcgcatctg	ccgcagcacc	ggatgcatct	gggaaaagca	aaaaaggggc	22020
tcgtccctgt	ttccggagga	atgttgcaagc	gggtccttgc	atgacgggga	ggcaaaccoc	22080
cgttcgcgcg	agtccggccg	gcccagagact	cgaaccgggg	gtcctgcgac	tcaacccttg	22140
gaaaataacc	ctccggctac	agggagcgag	ccacttaatg	ctttcgcttt	ccagcctaac	22200
cgcttacgcc	gcgcgcggcc	agtggccaaa	aaagctagcg	cagcagccgc	cgcgcctgga	22260
aggaagccaa	aaggagcgct	cccccggttg	ctgacgtcgc	acacctgggt	tcgacacgcg	22320
ggcggttaacc	gcatggatca	cggcggacgg	ccggatccgg	ggttcgaacc	ccggtcgtcc	22380
gccatgatac	ccttgcggaat	ttatccacca	gaccacggaa	gagtgcgccg	ttacaggctc	22440
tccttttgca	cggctctagag	cgtcaacgac	tgcgcacgcc	tcaccggcca	gagcgtcccg	22500
accatggagc	actttttgce	gctgcgcaac	atctggaacc	gcgtccgcga	ctttccgcgc	22560

gcctccacca	ccgccgcg	catcacctgg	atgtccaggt	acatctacgg	atatcatcgc	22620
cttatgttgg	aagacctcgc	ccccggagcc	ccggccaccc	tacgctggcc	cctctaccgc	22680
cagccgcgcg	cgcacttttt	ggtgggatat	cagtacctgg	tgccgacttg	caacgactac	22740
gtctttgact	caagggctta	ctcgcgtctc	aggtaacacc	agctctcgca	gccgggtcac	22800
cagaccgtta	actggtccgt	tatggccaac	tgcacttaca	ccatcaacac	gggcgcatac	22860
caccgctttg	tggacatgga	tgacttccag	tctacctca	cgcaggtgca	gcaggccata	22920
ttagccgagc	gcgttgctgc	cgacctggcc	ctgcttcagc	cgatgagggg	cttcgggggtc	22980
acacgcacatg	gaggaagagg	gcgccaccta	cggccaaact	ccgccgcgcg	cgtagcgata	23040
gatgcaagag	atgcaggaca	agaggaagga	gaagaagaag	tgccggtaga	aaggctcatg	23100
caagactact	acaaagacct	gcgccgatgt	caaaacgaag	cctggggcat	ggccgaccgc	23160
ctgcgcattc	agcaggccgg	acccaaggac	atggtgcttc	tgctcgacct	ccgcgctctc	23220
aagaccgcct	actttaatta	catcatcagc	agcaacctcg	ccagaaacaa	ccccgaccgc	23280
cacccgctgc	cgcccgccac	ggtgctcagc	ctaccttgcc	actgtgactg	gttagacgcc	23340
tttctcgaga	ggttttccga	tccggctgat	gcggactcgc	tcagggtccct	cggtggcgga	23400
gtacctacac	aacaattgtt	gagatgcac	gttagcgccg	tatccctgcc	gcacggcagc	23460
cccccgccaa	cccataaccg	ggacatgacg	ggcggcgctc	tccaactgcg	cccccgcgag	23520
aacgggcccgc	ccgtcaccga	gacctgcgcg	cgctcgccgc	gggagatgat	ccgcgctttt	23580
gtcgaccgcg	tcccgggtgc	ccgtcgtcgc	cgccgtgtcc	cccccccc	accgcgcgca	23640
gaagaagaag	aagaagggga	ggcccttatg	gaagaggaga	ttgaagaaga	agaggccctt	23700
gtagccctttg	agcgcgaggt	gcgcgacact	gtcgccgagc	tcattccgtct	tctggaggag	23760
gagttaaccg	tgtcggcgcg	caactcccag	tttttcaact	tcgccgtgga	cttctacgag	23820
gccatggagc	gccttgaggc	cttgggggat	atcaacgaat	ccacgttgcg	acgctggggt	23880
atgtactttt	tcgtggcgaga	acacaccgcc	accaccttca	actacctctt	tcagcgctcg	23940
cgaaactacg	ccgtcttcgc	ccggcacgtg	gagctcaatc	tcgcgcaggt	ggtcatgcgc	24000
gcccgcgatg	ccgaaggggg	cgtggtctac	agccgcgtct	ggaacgaggg	aggcctcaac	24060
gccttctcgc	agctcatggc	ccgcacctcc	aacgacctcg	ccgccaccgt	ggagcgagcc	24120
ggacgcggag	atctccagga	ggaagagatc	gagcagttca	tggccgaaat	cgccatataa	24180
gacaactcag	gagacgtgca	ggagattttg	cgccaggccg	ccgtcaacga	caccgaaatt	24240
gattctgtcg	aactctcttt	caggttcaag	ctcacccggc	ccgtcgtctt	ccgcagaggt	24300
cgccagattc	aggagatcaa	ccgcgcgctc	gtcgcgttgc	ccagcaacct	ccgcgcgcag	24360
caccagctcc	tgcccgcgcg	cggcgcgcag	gtgcccctgc	ccccctctcc	ggcggggtccc	24420
gagccccccc	tacctccggg	ggcccgcgcc	cgtcaccgct	tttagatgca	tcattccaagg	24480
acacccccgc	ggcccaccgc	ccgcgcgcgc	gtaccgtagt	cgccgcgcgg	ggatgcggcc	24540
tcttgcaagt	catcgacgcc	gccaccaacc	agcccctgga	aatcaggtat	cacctggacc	24600
tagccccgcg	cctgaccggc	ctatgcgagg	taaacctgca	ggagctcccg	cccgacctgt	24660
cgccgcggga	gctccagacc	atggacagct	cccatctgcg	cgatgttgct	atcaagctcc	24720
gaccgcgcgc	cgccgacatc	tggactttgg	gctcgcgcgc	cgtgggtggc	cgatccacca	24780
taactcccct	cgagcagcca	gacggtcaag	gacaagcagc	cgaagtagaa	gaccaccagc	24840
caaaccggcc	aggcgagggg	ctcaaattcc	cactctgctt	ccttgctgcg	ggtcgtcagg	24900
tcaacctcgt	gcaggatgta	cagcccgtgc	accgctgcca	gtactgcgca	cgttttttaca	24960
aaagccagca	cgagtgttcg	gcccgtcgca	gggacttcta	ctttcaccac	atcaacagcc	25020
actcctccaa	ctggtggcgg	gagatccagt	tcttcccgat	cggctcgcac	cctcgcaccg	25080
agcgtctctt	tgtcacctac	gatgtagaga	cctatacttg	gatggggggc	tttgggaagc	25140
agctcgtgcc	cttcatgctg	gttatgaagt	tcggcggaga	tgagcctctg	gtgaccgccc	25200
cgcgagacct	agccgtggac	cttggatggg	accgctggga	acaagaccgc	cttaccttct	25260
actgcacac	cccagaaaaa	atggccatag	gtcgcaggtt	taggaccttt	cgcgaccacc	25320
tgcaaatgct	aatggcccgt	gaactgtgga	gctcattcgt	cgcttccaac	cctcatcttg	25380
cagactgggc	cctgtcagaa	cacgggctca	gctccccctga	ggagctcacc	tacgaggaac	25440
ttaaaaaatt	gccctccatc	aagggcaccc	cgcgcttctt	ggaactttac	atcgtggggc	25500
acaacatcaa	cggcttcgac	gagatcgtgc	tcgccgccca	ggtaattaac	aaccgttccg	25560
aggtgcccgg	acccttccgc	atcacacgca	acttttatgcc	tcgcgcggga	aagatacttt	25620
tcaacgatgt	caccttcgcg	ctgccaaacc	cgcgttccaa	aaagcgcacg	gactttttgc	25680
tctggggagca	gggcggatgc	gacgacactg	acttcaaata	ccagtacctc	aaagtcatgg	25740
ttagggacac	ctttgcgctc	acccacacct	cgctccggaa	ggccgcgcag	gcatacgcgc	25800
taccgtaga	aaagggatgc	tgcgccctacc	aggccgtcaa	ccagttctac	atgctaggct	25860
cttaccgttc	ggaggccgac	gggtttccga	tccaagagta	ctggaaagac	cgcgaaagag	25920
ttgtcctcaa	ccgcgagctg	tggaaaaaaa	agggacagga	taagtatgac	atcatcaagg	25980
aaaccctgga	ctactgcgcg	ctagacgtgc	aggtcaccgc	cgagctggtc	aacaagctgc	26040

gagactccta	cgcttccttc	gtgcgtgacg	cggtaggtct	cacagacgcc	agottcaacg	26100
tcttccagcg	tccaaccata	tcatccaact	cacatgccat	cttcaggcag	atagtcttcc	26160
gagcagagca	gcccgcctcg	agcaacctcg	gtcccgcact	cctcgctccc	tgcacgaa	26220
tatacgatta	cgtgcgcgcc	agcatccgcg	gtggaagatg	ctaccctaca	tatcttggaa	26280
tactcagaga	gccccctctac	gtttacgaca	tttgcggtat	gtacgcctcc	gcgtcaccc	26340
accccatgcc	atgggggtccc	ccactcaaac	catacgagcg	cgcgcttgcc	gcccgcgcat	26400
ggcagcaggg	gctagacttg	caaggatgca	agatagacta	cttcgacgcg	cgctgctgc	26460
ccgggggtctt	taccgtggac	gcagaccccc	cggacgagac	gcagctagac	ccactaccgc	26520
cattctgttc	gcgcaagggc	ggcgcgcctct	gctggaccaa	cgagcgccta	cgcgagaggg	26580
tagccaccag	cgttgacctt	gtcacccctgc	acaaccgcgg	ttggcgcgtg	cacctggtgc	26640
ccgacgagcg	caccaccgtc	tttcccgaat	ggcgggtgcgt	tgcgcgcgaa	tacgtgcagc	26700
taaacatcgc	ggccaaggag	cgcgccgatac	gcgacaaaaa	ccaaacctg	cgctccatcg	26760
ccaagttgct	gtccaacgcc	ctctacgggt	cgtttgccac	caagcttgac	aacaaaaaga	26820
ttgtcttttc	tgaccagatg	gacgcggcca	ccctcaaagg	catcaccgcg	ggccaggtga	26880
atatcaaatac	ctcctcgttt	ttggaaactg	acaatcttag	cgcagaagtc	atgcccgcctt	26940
ttgagagggga	gtactcaccc	caacagctgg	ccctcgacga	cagcgatgcg	gaagagagtg	27000
aggacgaacg	gcccccacc	cccttttata	gcccccttc	aggaacaccc	ggctgctgg	27060
cctacaccta	taaaccaatac	accttccttg	gtcccgaaga	ggcgacatg	tgtcttcaca	27120
ccctggagcg	agtggacccc	ctagtggaca	acgaccgcta	cccctcccac	ttagcctcct	27180
tcgtgctggc	ctggacgcga	gccttcgtct	cagagtggtc	cgagtttcta	tacgaggagg	27240
accgcggaac	accgctcgag	gacaggcctc	tcaagtctgt	atacggggac	acggacagcc	27300
ttttcgtcac	cgagcgtgga	caccggctca	tggaaaccag	aggtaagaaa	cgcatcaaaa	27360
agcatggggg	aaacctgggt	tttgaccocg	aacggccaga	gctcacctgg	ctcgtggaat	27420
gcgagaccgt	ctgcggggcc	tgccggcgccg	atgcctactc	cccggaatcg	gtatttctcg	27480
cgcccaagct	ctacgccctt	aaaagtctgc	actgcccttc	gtgcggcgcc	tcctccaagg	27540
gcaagctgcg	cgccaagggc	cacgcgcgcg	aggggctgga	ctatgacacc	atgggtcaaat	27600
gctacctggc	cgacgcgcag	ggcgaagacc	ggcagcgctt	cagcaccagc	aggaccagcc	27660
tcaagcgcac	cctggccagc	gcgcagcccg	gagcgcaccc	cttcaccgtg	accagacta	27720
cgctgacgag	gaccctgcgc	ccgtggaaag	acatgacctt	ggcccgctctg	gacgagcacc	27780
gactactgcc	gtacagcgaa	agccgcacca	acccgcgaaa	cgaggagata	tgctggatcg	27840
agatgccgta	gagcacgtga	ccgagctgtg	ggaccgcctg	gaactgcttg	gtcaaacgct	27900
caaaagcatg	cctacggcgg	acggcctcaa	accgttgaaa	aactttgctt	ccttgcaaga	27960
actgctatcg	ctggggcggc	agcgccttct	ggcgcatttg	gtcagggaaa	acatgcaagt	28020
cagggacatg	cttaacgaag	tggccccctt	gctcagggat	gacggcagct	gcagctctct	28080
taactaccag	ttgcagccgg	taataggtgt	gatttacggg	cccaccggct	gcggtaagtc	28140
gcagctgctc	aggaacctgc	tttcttccca	gctgatctcc	cctaccccg	aaacgggtttt	28200
cttcacgcgc	ccgcaggtag	acatgatccc	cccatctgaa	ctcaaagcgt	gggaaatgca	28260
aatctgtgag	ggtaactacg	cccctgggccc	ggatggaaac	attataccgc	agtctggcac	28320
cctccgccc	cgctttgtaa	aaatggccta	tgacgatctc	atcctggaa	acaactatga	28380
cgtagtgat	cccagaaata	tcttcgcccc	ggcgcgcgcc	cgtgggcccc	ttgccatcat	28440
tatggacgaa	tgcatggaaa	atctcgagg	tcacaagggc	gtctccaagt	tcttccacgc	28500
atctccttct	aagctacatg	acaaatttcc	caagtgcacc	ggatacactg	tgctgggtggt	28560
tctgcacaac	atgaatcccc	ggagggatat	ggctgggaac	atagccaacc	taaaaaataca	28620
gtccaagatg	catctcatat	ccccacgtat	gcacccatcc	cagcttaacc	gctttgtaaa	28680
cacttacacc	aagggcctgc	ccctggcaat	cagcttgcta	ctgaaagaca	tttttaggca	28740
ccacgcccag	cgctcctgct	acgactggat	catctacaac	accaccccg	agcatgaagc	28800
tctgcagtgg	tgctacctcc	accccagaga	cgggcttatg	cccatgtatc	tgaacatcca	28860
gagtcacctt	taccacgtcc	tggaaaaaat	acacaggacc	ctcaacgacc	gagaccgctg	28920
gtcccgggcc	taccgcgcgc	gcaaaacccc	taaataaaga	cagcaagaca	cttgcttgat	28980
caaaatccaa	acagagtctg	gttttttattt	atgtttttaa	ccgcattggg	aggggaggaa	29040
gccttcaggg	cagaaacctg	ctggcgcgaga	tccaacagct	gctgagaaa	gacattaagt	29100
tcccgggtca	aagaatccaa	ttgtgccaaa	agagccgtca	acttgctcatc	gcgggcggat	29160
gaacgggaag	ctgcactgct	tgcaagcggg	ctcaggaaaag	caaagtcagt	cacaatccc	29220
cgggcggtgg	ctgcagcggc	tgaagcggcg	gcggaggctg	cagtctccaa	cggcgttcca	29280
gacacggtct	cgtaggtcaa	ggtagtagag	tttgccgggca	ggacggggcg	accatcaatg	29340
ctggagccca	tcacattctg	acgcaccccc	gcccattggg	gcatgcgcgt	tgtaaaatat	29400
gagctcacia	tgcttccatc	aaacgagttg	gcgctcatgg	cggcggtgc	tgcaaaacag	29460
atacaaaact	acatgagacc	cccaccttat	atattctttc	ccacccttaa	gccccgcccc	29520

tcgatggcaa	acagctatta	tgggtattat	gggtgctagc	gacatgaggt	tgccccgtat	29580
tcagtgtcgc	tgatttgtat	tgtctgaagt	tgtttttacg	ttaagttgat	gcagatcaat	29640
taatacgata	cctgcgtcat	aattgattat	ttgacgtggt	ttgatggcct	ccacgcacgt	29700
tgtgatatgt	agatgataat	cattatcact	ttacgggtcc	tttccggtga	tccgacaggt	29760
tacggggcgg	cgacctcgcg	ggttttcgct	atttatgaaa	attttccggt	ttaaggcggt	29820
tccgttcttc	ttcgtcataa	cttaatgttt	ttatttaaaa	tacctcttga	aaagaaagga	29880
aacgacaggt	gctgaaagcg	aggctttttg	gcctctgtcg	tttcttttct	ctgtttttgt	29940
ccgtggaatg	aacaatggaa	gttaacggat	ccaggccgcg	agcaaaaagg	cagcaaaaagg	30000
ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	30060
agcatcacaa	aatcaacgc	tcaagtcaga	gggtggcga	cccgcacagga	ctataaagat	30120
accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	30180
ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	30240
gtaggtatct	cagttcgggtg	taggtcgttc	gctccaagct	gggctgtgtg	cacgaacccc	30300
ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	aaccgggtaa	30360
gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggtatg	30420
taggcgggtg	tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	agaagaacag	30480
tatttgggtat	ctgcgctctg	ccaaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	30540
gatccggcaa	acaaaccacc	gctggtagcg	gtgggttttt	tgtttgcaag	cagcagatta	30600
cgcgagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	30660
agtggaaacga	aaactcacgt	taagggatct	tgggtcatcag	attatcaaaa	aggatcttca	30720
cctagatcct	tttaaattaa	aatgaagtt	ttaaataaat	ctaaagtata	tatgagtaaa	30780
cttgggtctga	cagttacca	tgcttaata	gtgaggcacc	tatctcagcg	atctgtctat	30840
ttcgttcatc	catagttgcc	tgactccccg	tagtgtagat	aactacgata	cgggaggggt	30900
taccatccgg	ccccagtgtc	gcaatgatac	cgctgaccc	acgctcaccg	gctcctgatt	30960
tatcagcaat	aaaccagcca	gccggaagtg	ccgagcgcag	aagtggctct	gcaactttat	31020
ccgcctccat	ccagtctatt	agttgttgcc	gggaagctag	agtaagtagt	tcgccagtta	31080
atagttttcg	caacgttggt	gccattgcta	caggcatcgt	ggtgtcacgc	tcgtcgtttg	31140
gstatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	tcccccatgt	31200
tgtgcaaaaa	agcggttagc	tccttcggtc	ctccgatagt	tgtcagaagt	aagttggccg	31260
cagtgttatc	actcatgggt	atggcagcac	tgcataattc	tcttactgtc	atgccatccg	31320
taagatgctt	ttctgtgact	ggtgagtatt	caaccaagaa	tacgggataa	taccgcgcca	31380
catagcagaa	ctttaaaagt	gctcatcatt	gggaaacggt	cttcggggcg	aaaactctca	31440
aggatcttac	cgctgttgag	atccagttcg	atgtaaccca	ctcgcgcacc	caagtgatct	31500
tctgcatctt	ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaatgcc	31560
gcaaaaaagg	gaataagggc	gacacggaaa	tggtgaatac	tcatactttt	cctttttcaa	31620
tattattgaa	gcatttatca	gggttattgt	ctcatcagcg	gatacatatt	tg	31672

<210> 4

<211> 30365

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 4

gaatcggcca	gcgcaatta	actataacgg	tcctaaggta	gcgtcatcat	cataatatac	60
cttatttttg	attgaagcca	atatgataat	gaggggggtg	agtttgtgac	gtggcgcggg	120
gcgtgggaac	ggggcggggtg	acgtaggttt	tagggcggag	taacttgcat	gtattgggaa	180
ttgtagtttt	tttaaaatgg	gaagttacgt	atcgtgggaa	aacggaagtg	aagatttgag	240
gaagttgtgg	gttttttggc	tttcgtttct	gggcgtaggt	tcgcgtgcgg	ttttctgggt	300
gttttttgtg	gactttaacc	gttacgtcat	tttttagtcc	tatatatact	cgctctgtac	360
ttggcccttt	ttacactgtg	actgattgag	ctggtgcccgt	gtcgagtggg	gttttttaaat	420
agggtttttt	actggtaagg	ctgactgtta	tggctgcccgc	tgtggaagcg	ctgtatgttg	480
ttctggagcg	ggagggtgct	attttgccta	ggataaactc	gtataatgta	tgctatacga	540
agttatggcg	cgccagatct	gtttgtcacg	cccgcacctg	gttttgcttc	aggaaatatg	600

actacgtccg	gcgttccatt	tggcatgaca	ctacgaccaa	cacgatctcg	gttgtctcgg	660
cgcactccgt	acagtaggga	tcgcctacct	ccttttgaga	cagagacccg	cgctaccata	720
ctggaggatc	atccgctgct	gcccgaatgt	aacactttga	caatgcacaa	cgtaggattac	780
gtgcgaggtc	ttccctgcag	tgtgggattt	acgctgattc	aggaatgggt	tgttccctgg	840
gatattggtc	tgacgcggga	ggagcttgta	atcctgagga	agtgtatgca	cgtgtgcctg	900
tgtttgtcca	acattgatat	catgacgagc	atgatgatcc	atggttacga	gtcctgggct	960
ctccactgtc	attgttccag	tcccggttcc	ctgcagtgca	tagccggcgg	gcagggttttg	1020
gccagctggt	ttaggatggt	ggtggatggc	gccatgttta	atcagagggt	tatatggtac	1080
cgggaggtgg	tgaattacaa	catgccaaaa	gaggtaatgt	ttatgtccag	cgtgtttatg	1140
aggggtcgcc	acttaatcta	cctgcgcttg	tggtatgatg	gccacgtggg	ttctgtggtc	1200
cccgccatga	gctttggata	cagcgcttg	cactgtggga	ttttgaacaa	tattgtgggtg	1260
ctgtgctgca	gttactgtgc	tgatttaagt	gagatcaggg	tgcgctgctg	tgcccggagg	1320
acaaggcgtc	tcattgctgcg	ggcgggtgcga	atcatcgctg	aggagaccac	tgccatgttg	1380
tattcctgca	ggacggagcg	gcggcggcag	cagtttatct	gcgcgctgct	gcagcaccac	1440
cgccttatcc	tgatgcacga	ttatgactct	acccccatgt	aggcgtggac	ttccctctcg	1500
ccgcccgttg	agcaaccgca	agttggacag	cagcctgtgg	ctcagcagct	ggacagcgac	1560
atgaacttaa	cgcagctgcc	cggggagttt	attaatatca	ctgatgagcg	tttggctcga	1620
caggaaacccg	tgtggaatat	aacacctaa	aatatgtctg	ttacccatga	tatgatgctt	1680
tttaaggcca	gccggggaga	aaggactgtg	tactctgtgt	gttgggaggg	aggtggcagg	1740
ttgaatacta	gggttctgtg	agtttgatta	aggtagcgtg	atcaatataa	gctatgtggt	1800
ggtggggcta	tactactgaa	tgaaaaatga	cttgaaattt	tctgcaattg	aaaaataaac	1860
acgttgaaac	ataacatgca	acaggttcac	gattctttat	tcttgggcaa	tgtaggagaa	1920
gggtgaagag	ttggtagcaa	aagtttcagt	gggtgatctt	ccactttccc	aggaccatgt	1980
aaaagacata	gagtaagtgc	ttacctcgct	agtttctgtg	gattcactag	tgccattaag	2040
tgtaatggta	agtatcatag	gttttagttt	atcaccatgc	aagtaaactt	gactgacaat	2100
gttattttta	gcagtttgac	tttgggtttt	tggataggct	agaaggtag	gcataaatcc	2160
aactgcattt	gtgtatggat	ttgcattagt	tgagttccca	tttctaaagt	tccagtaatg	2220
tttttttaag	gaggagtctt	ccattagaac	accgttttgg	tcaaactctaa	ggaatatact	2280
aacacttgca	acggtgcctg	tcattggatga	aagatctcca	gatacagcca	aagcagctac	2340
agtagctagt	acttgactcc	cacattttgt	aagaacccaa	gtaaatttgc	agtcattatc	2400
tgaatgaatt	ctgcagttag	gagatgggtc	tggggttgct	cacagggtaa	gtttgtcatc	2460
atttttgttt	cctattgtaa	tggccctga	gttgtcaaa	cttaaacccg	ctccaagttt	2520
agtaatcatg	gcaccgtttt	cattgtaatc	aatgccagag	ccaattttag	tttttatttg	2580
gttgatatct	ggagactcag	atgtgtttgt	atcaaaactc	agacccttcc	ctgcatttat	2640
agctatggca	gtattatcaa	agtttagtcc	actggatttt	tttatgctaa	cttccagttt	2700
tttagtattg	tttgatgcat	taaaaaggta	taggcctctg	ttatagttta	tgtccaagtt	2760
atgagatgca	ttaatataca	ggggtccctg	cccagttta	agacgtagtt	ttgtttgagc	2820
atcaaatggg	taatccacat	ctagaattaa	caagtgtgta	tttatacgca	tgccaccgcc	2880
cgttttaatt	tccatgttgt	ttgatgaatc	ataaccaata	gctcctgcaa	ctttggttct	2940
aaggggagttt	tgttcaacgg	tgacacctgg	tccagtaact	actgttagtg	tactcggagt	3000
ttgtgctact	tgcaaaggac	cgttattttt	aattcctatt	tttccattat	ttacataaat	3060
aggatcttcc	atgttaatgc	ccaagctacc	cgtggcagta	gttagcgggg	gtgatgcagt	3120
tacagtaagg	gtgtcgctgt	cactgccaga	gaggggggct	gatgtttgca	gggctagctt	3180
tccatctgac	actgtaatgg	gcccttttag	agcaatgctt	agtttggagt	cttgcacggg	3240
cagtggggct	tgtgactgta	cgctaagagc	gccgctagta	actatcagag	gagcgggtgt	3300
tgccactgtt	agggcgcttg	aggtaattgt	aagtgggtcg	gaggtgtcca	aacttatgtt	3360
tgactttgtt	tttttaagtg	gctgagtaac	agtgggttaca	ttttgggagg	tgagggtttcc	3420
ggccttgtct	agggtaagac	cgctgcccat	tttaagcgca	agcatgccgt	gggaggtgtc	3480
caaagggttcg	gagacgcgta	gagagagaa	tccaggggga	ctttcttggga	aaccattggg	3540
tgaaacaaat	ggaggggtaa	gaaagggcac	agttggaggc	cgggtttctg	tgtcatatgg	3600
atacacgggg	ttgaagggtg	cttcagacgg	tctggcgcgt	ttcatctgca	acaatatgaa	3660
gatagtgggt	gcggagggac	aagaacatga	ggaatttgac	atcccattta	aactttggag	3720
aaagtgttga	gctaaaaggc	ggctgagata	ccagagttag	gaggaaggaa	aggaggtgat	3780
gctgaataag	ctggacaaag	atgtgtgac	tgatttttaag	taagtaattt	attcagtcgt	3840
agccgtccgc	cgagtctttc	accgcgtcaa	agttgggaat	aaactgggtc	gggtagtggc	3900
cgggaggtcc	agaaaagggg	ttgaagtaaa	ccgaaggcac	gaactcctca	ataaattgta	3960
gagttccaat	gcctccggag	cgcggctccg	aggacgaggt	ctgcagagtt	aggatcgctt	4020
gacggggcgt	aaatgaagag	cggccagcgc	cgcgatctg	aatgtcccg	tccggacgga	4080

gaccaagaga	ggagctcacc	gactcgtcgt	tgagctgaat	acctcgcct	ctgattttca	4140
ggtgagttat	accctgccc	ggcgaccgca	ccctgtgacg	aaagccgccc	gcaagctgcg	4200
cccctgagtt	agtcatctga	acttcggcct	gggctgtctt	gggaagtacc	acagtggtgg	4260
gagcgggact	ttcctggtac	accagggcag	cgggccaact	acggggatta	aggttattac	4320
gaggtgtggt	ggtaatagcc	gcctgttcga	ggagaattcg	gtttcgggtg	gcgcggattc	4380
cgttgacccg	ggatatcatg	tgggggtccc	cgctcatgta	gtttattcgg	gttgagtagt	4440
cttgggcagc	tccagccgca	agtcccat	gtggctggta	actccacatg	tagggcgtgg	4500
gaatttcctt	gctcataatg	gcgctgacga	caggtgctgg	cgccgggtgt	ggccgctgga	4560
gatgacgtag	ttttcgcgct	taaatttgag	aaagggcgcg	aaactagtcc	ttaagagtca	4620
gcgcgcagta	tttgctgaag	agagcctccg	cgtcttcacg	cgtgcgccga	agctgatctt	4680
cgtttttgtg	atacaggcag	ctgcgggtga	gggagcgca	agacctgttt	tttattttca	4740
gctcttggtt	ttggcccttg	ctttgttgaa	atatagcata	cagagtggga	aaaatcctat	4800
ttctaagctc	gcgggtcgat	acgggttcgt	tgggcgccag	acgcagcgct	cctcctcctg	4860
ctgctgcgcg	cgctgtggat	ttcttgggct	ttgtcagagt	cttgctatcc	ggtcgccttt	4920
gcttctgtgt	gaccgctgct	gttgcctgcg	ctgccgctgc	cgccggtgca	gtaggggctg	4980
tagagatgac	ggtagtaatg	caggatgtta	cgggggaagg	ccacgccgtg	atggtagaga	5040
agaaagcggc	gggcgaagga	gatgttgc	ccactgtctt	gcaagcaagc	aactatggcg	5100
ttcttgtgcc	cgccacga	gcggtagcct	tggcgctgtt	gttgctcttg	ggctaaccgc	5160
ggcggctgct	tagacttacc	ggccttggtt	ccagtgggtg	cccactctacg	gttgggtcgg	5220
cgaacaggca	gtgccggcgg	cgccctgagga	gcggagggtg	tagcgatgct	gggaacgggt	5280
gccaatttct	ggggcgccgg	cgaggggaat	gcgaccgagg	gtgacgggtg	ttcgtctgac	5340
acctcttcgg	cctcggaagc	ttcgtctagg	ctgtcccagt	cttccatcat	ctcctcctcc	5400
tgcgtcaaaa	cctcctctgc	ctgactgtcc	cagtattcct	cctcgtccgt	gggtggcgcc	5460
ggcggcagct	gcagcttctt	tttgggtg	atcctgggaa	gcaagggccc	gcggctgctg	5520
atagggctgc	ggcggcgggg	ggattgggtt	gagctcctcg	ccggactggg	ggtccaggta	5580
aacccccctg	ccctttcgtg	gcagaaactc	ttggcgggct	ttggtgatgg	cttgcaattg	5640
gccaaggatg	tggccctggg	taatgacgca	ggcggtaagc	tccgcatttg	gcgggcggga	5700
ttggtcttcg	tagaacctaa	tctcgtgggc	gtggtagtcc	tcaggtacaa	atttgcgaag	5760
gtaagccgac	gtccacagcc	ccggagttag	tttcaacccc	ggagccgcgg	acttttcgtc	5820
aggcgaggga	ccctgcagct	caaaggatcc	gataatttga	ctttcgctaa	gcagttgcga	5880
attgcagacc	agggagcggg	gcgggggtgca	taggttgtag	cgacagttag	actccagttag	5940
gccgtcaccg	ctcacgtctt	ccatgatgtc	ggagtggtag	gcaaggtagt	tggttagctg	6000
cagaaggtag	cagtacccc	aaagcggcgg	agggcattca	cggtacttaa	tgggcacaaa	6060
gtcgtctagga	agcgcacagc	aggtggcggg	cagaattcct	gaacgctcta	ggataaagtt	6120
cctaaagtgt	tgcaacatgc	tttgactggg	gaagtctggc	agacctgtt	agaggggttt	6180
aagcaggcgt	tcggggaaga	taatgtccgc	caggtgcgcg	gccacggagc	gctcgttgaa	6240
ggcgtccat	aggtccttca	agttttgctt	tagcagcttc	tgagctcct	ttaggttgcg	6300
ctcctccagg	cattgctgcc	acacgccc	ggcgtttg	caggtgtagc	acagaaataa	6360
gtaaacgcag	tcgcggacgt	agtcgcggcg	cgcctcgc	ttgagcgtgg	aatgaagcac	6420
gttttgcccg	aggcggtttt	cgtgcaaaat	tccaggtag	gagaccaggt	tgacagctc	6480
cacgttgga	attttgacag	cctggcgac	gtagccctgg	cgaaagggtg	actgcaacgt	6540
ttcctctagc	ttgcgctgca	tctccgggtc	agcaaaagaa	cgctgcatgc	actcaagctc	6600
cacggtaaca	agcactgcgg	ccatcattag	cttgctgcg	tcctccaagt	cgccaggctc	6660
gcgcgtctca	agccagcgcg	ccagctgctc	atcgccaa	gcgggtaggc	cctcctcggt	6720
ttgttcttgc	aagtttgcat	ccctctccag	gggtcgtgca	cgccgcacga	tcagctcgct	6780
catgactgtg	ctcataacct	tggggggtag	gttaagtggc	gggtaggcaa	agtggttgac	6840
ctcgatgctg	cgtttcagca	cgctaggcg	cggttggtca	ccctcaagtt	ccaccagcac	6900
tccacagtga	ctttcat	cgctgttttc	ttgttgacga	gcgtttgccc	cgcggtttct	6960
gtcgcgtcca	agaccctcaa	agatttttgg	cacttcgtcg	agcgaggcga	tatcaggtag	7020
gacagcgccc	tgccgcaagg	ccagctgctt	gtccgctcgg	ctgcggttgg	cacggcagga	7080
taggggtatc	ttgcagtttt	ggaaaaagat	gtgatagggtg	gcaagcacct	ctggcacggc	7140
aaatacgggg	tagaagttga	ggcgcgggtt	gggctcgcat	gtgccgtttt	cttggcggtt	7200
gggggggtacg	cgcggtgaga	acaggtggcg	ttcgtaggca	aggctgacat	ccgctatggc	7260
gaggggcaca	tcgctgcgct	cttgcaacgc	gtcgcagata	atggcgccact	ggcgtgcag	7320
atgcttcaac	agcacgtcgt	ctccacatc	taggtagtgc	ccatgccttt	ggtccccccg	7380
cccgacttgt	tcctcgtttg	cctctgcgtc	gtcctggtct	tgctttttat	cctctgttgg	7440
tactgagcga	tcctcgctcg	cttcgcttac	aaaacctggg	tcctgctcga	taatcacttc	7500
ctcctcctca	agcgggggtg	cctcgacggg	gaaggtggtg	ggcgcgttgg	cggcatcggt	7560

ggaggcgggtg	gtggcgaaact	caaagggggc	ggttaggctg	tcctccttct	cgactgactc	7620
catgatcttt	ttctgcctat	aggagaagga	aatggccagt	cgggaagagg	agcagcgcg	7680
aaccaccccc	gagcgcggac	gcggtgcggc	gcgacgtcca	ccaaccatgg	aggacgtgtc	7740
gtccccgtcg	ccgtcgcgcg	cgccctcccc	cgcgccccca	aaaaagcggc	tgaggcggcg	7800
tctcgagtcc	gaggacgaag	aagactcgtc	acaagatgcg	ctggtgcgcg	gcacaccag	7860
cccgcggcca	tcgacctcga	cggcggatgt	ggccattgcg	tcacaaaaga	aaaagaagcg	7920
cccctctccc	aagcccagac	gcccgcctac	cccagagggtg	atcgtggaca	gcgaggaaga	7980
aagagaagat	gtggcgctac	aaatggtggg	tttcagcaac	ccaccggtgc	taatcaagca	8040
cggcaaggga	ggtaagcgca	cgggtgcggcg	gctgaatgaa	gacgaccag	tggcgcgggg	8100
tatgcggacg	caagaggaaa	aggaagagtc	cagtgaagcg	gaaagtga	gcacgggtgat	8160
aaacccgctg	agcctgccga	tcgtgtctgc	gtgggagaag	ggcatggagg	ctgcgcgcgc	8220
gttgatggac	aagtaccacg	tggataacga	tctaaggca	aacttcaagc	tactgcctga	8280
ccaagtggaa	gctctggcgg	ccgtatgcaa	gacctggcta	aacgaggagc	accgcggggtt	8340
gcagctgacc	ttcaccagca	acaagacctt	tgtgacgatg	atggggcgat	tcctgcaggc	8400
gtacctgcag	tcgtttgcag	aggtaaccta	caagcaccac	gagcccacgg	gctgcgcggt	8460
gtggctgcac	cgctgcgctg	agatcgaagg	cgagcttaag	tgtctacacg	ggagcattat	8520
gataaataag	gagcacgtga	ttgaaatgga	gtgacagc	gaaaacgggc	agcgcgcgct	8580
gaaggagcag	ctagcaagg	ccaagatcgt	gagaacccg	tggggccgaa	atgtggtgca	8640
gatctccaac	accgacgcaa	ggtgctgcgt	gcatgacgcg	gcctgtccgg	ccaatcagtt	8700
ttccggcaag	tcttgccgca	tgttcttctc	tgaaggcgca	aaggctcagg	tggcttttaa	8760
gcagatcaag	gctttcatgc	aggcgtgtga	tcctaaccgc	cagaccgggc	acggtcacct	8820
tctgatgcca	ctacggtgcg	agtgcacctc	aaagcctggg	catgcacctt	ttttgggaag	8880
gcagctacca	aagttgactc	cgcttcgcct	gagcaaccgc	gaggacctgg	acgcggatct	8940
gatctccgac	aagagcgtgc	tggccagcgt	gcaccaccgc	gcgctgatag	tgttccagtg	9000
ctgcaaccct	gtgtatcgca	actcgcgcgc	gcagggcgga	ggccccaact	gcgacttcaa	9060
gatatcgggc	cccgcacctgc	taaaccgcgt	ggtgatgggtg	cgagcctgt	ggagtga	9120
cttcaccgag	ctgcgcggga	tggttgtgct	tgagttaag	tggagcacta	aacaccagta	9180
tcgcaacgtg	tccttgccag	tggcgcatag	cgatgcgcgg	cagaaccctt	ttgattttta	9240
aacggcgctg	acggcaagg	tggggggtaa	ataatcacc	gagagtgtac	aaataaaaac	9300
atctgctctt	attgaaagt	tctcctagta	cattattttt	acatgttttt	caagtgacaa	9360
aaagaagtgg	cgctccta	ctgcgcactg	tggctgcgga	agtagggcga	gtggcgctcc	9420
aggaagctgt	agagctgttc	ctggttgcga	cgcaggggtg	gctgtacctg	gggactgtta	9480
agcatggagt	tgggtacccc	ggtaataagg	ttcatgggtg	ggttgtgatc	catgggagtt	9540
tggggccag	tggcaaaggc	gtggagaaac	atgcagcaga	atagtccaca	ggcgcccgag	9600
ttgggcccct	gcacgctttg	ggtggacttt	tcacagctta	tacagcggtc	gggggaagaa	9660
gcaatggcgc	tacggcgcag	gagtgaactg	tactcaaact	ggtaaacctg	cttgagtcgt	9720
tggtcagaaa	agccaaagg	ctcaaagagg	tagcatgttt	ttgagcgcgg	gttccaggca	9780
aaggccatcc	agtgtacgcc	cccagtcctg	cgaccggccg	tattgactat	ggcgagggcg	9840
agcttgtgtg	gagaaacaaa	gcctggaaag	cgcttgtcat	aggtgcccaa	aaaatatggc	9900
ccacaaccaa	gatctttgac	aatggccttt	agttcctgct	cactggagcc	catggcgcca	9960
gctgttggtg	atgttgcttg	cttcttttat	gttgtggcgt	tgccggccga	gaagggcggtg	10020
cgcagggtaca	cggctctcgat	gacgcgcggg	tgccgctggt	gcacacggac	cacgtcaaag	10080
acttcaaaca	aaacataaag	aaggggtggg	tcgtccatgg	gatccacctc	aaaagtcag	10140
tctagcgcgt	gggaggagtt	ggcgtagaga	aggttttggc	ccaggtctgt	gagtgcgccc	10200
atggacataa	agttactgga	gaatgggatg	cgccaaagg	tgcgatcgca	aagaaacttt	10260
ttctgggtaa	tactgtcaac	cgcggttttg	cctattagtg	ggtagggcac	gttggcgggg	10320
taagcctgtc	cctcgcgcac	ggtgggagcg	aggtagccta	cgaatcctga	gttggtatgc	10380
tgggtgaagaa	ttccaacctg	ctgatactcc	ttgtatttag	tatcgtcaac	cacttgccgg	10440
ctcatgggct	ggaagtctct	gaagaacgag	tacatgcggt	ccttgtagct	ttctggaatg	10500
tagaagccct	ggtagccaat	attgtagttg	gccaacatct	gcaccaggaa	ccagtccttg	10560
gtcatgttgc	actgagctac	gttgtagccc	tccccgtcaa	ctgagcggtt	aatctcaaac	10620
tcattgggag	taagcaggcg	gtcgttgccc	ggccagctaa	cagaagagtc	aaaggtaatg	10680
gccaccttct	taaagggtgtg	attaagatag	aagggtccgt	caaggatatg	tatggagcca	10740
gagtaggtgt	agtaagggtc	gtagcctgat	cccagggaag	gggtttcctt	tgtcttcaag	10800
cgtgtgaagg	cccaaccgcg	aaatgctgcc	cagttgcgcg	atgggatgga	gatgggcacg	10860
ttggtggcgt	tggcgggtat	ggggtatagc	atggtggcgg	cggaaaggta	gtcattaaag	10920
gactggtcgt	tgggtgtcatt	tctgagcatg	gcttccagcg	tggaggccgt	gttgtgggcc	10980
atggggaaga	aggtggcgta	aagacaaatg	ctgtcaaac	taatgctagc	cccgtaact	11040

ctaagatcgt	ttcccagaga	gctctgcaga	accatgttaa	catccttcct	gaagttccat	11100
tcatatgtat	atgagcctgg	caggaggagg	aggttttttaa	tggcaaaaaa	cttttggggc	11160
acctgaatgt	gaaagggcac	gtagcggcgc	tttcccaaca	acatggagcg	ataacggagg	11220
cccgcatgtc	ggtggtggtt	aaagggatta	acgttgtcca	tgtagtccag	agaccagcgc	11280
gccccagggt	taatgtagca	gtctacaagc	ccgggagcca	ccactcgctt	gttcattgtag	11340
tcgtaggtgt	tgggggttgc	agatatttcc	acattggtgg	ggttgtatct	tagcttgtct	11400
ggcaggtaca	gcgcaatatt	ggagtaaagg	aaatttctcc	ataggttggc	atthaggtta	11460
atttccatgg	caaagttggt	accactcct	atttcattac	gtgttgcaaa	agtttcatct	11520
tttgtccatg	tagtatctcc	attatcgct	gagccattgc	cattagcctt	aatagcttga	11580
taggtgtcag	ttaccccaat	acccccaaga	ggaaaacaat	aatttggcaa	ttcatcctca	11640
gttccatggt	tttcaatgat	tctaacatct	ggatcatagc	tgtctacagc	ctgattccac	11700
atagaaaaat	atctggttct	atcacctatg	gaatcaagca	agagttgata	ggacagctct	11760
gtgtttctgt	cttgcaaatc	taccacggca	tttagctgcg	atgcctgacc	agcaagaaca	11820
cccatgttgc	cagtgcgtgt	ataatacatt	aggccaataa	aattgtccct	gaaagcaatg	11880
taattgggtc	tgtttggcat	agattgttga	cccaacatag	ctttagaatt	ttcatcacct	11940
tttccagggt	tgtaagacag	atgtgtgtct	ggggtttcca	tatttatact	ttcattgtac	12000
aaaaccactt	ttggtttagt	agcattgcct	tggcgtcgt	tcaaagaggt	agtatttgag	12060
aagaattgca	agtcaacctt	tgggaagggc	accccttttt	catccggaac	cagaacggat	12120
tgaccaccaa	aaggatttgt	aggcctggca	taagatccat	agcatggttt	catgggagtt	12180
gttttttttaa	gcactctccc	tcctgcgcga	ttagcatcag	cttcggtcca	ctgagattcg	12240
ccaatttgag	gttctgggtg	ataggaagga	tctgcgtata	cagggttagc	ttgtgtttct	12300
gcattgtctg	atcctatttg	tagccgcctt	tttgtaatgt	tttctccaga	caaaggagcc	12360
tgggcataga	catgtgtttt	cttagtagcc	tgatctcgag	cgttttgtct	ttcttcttcc	12420
tcttcttcat	cttcatcttc	ctcttcttca	tcctcggcaa	ctgcccggcc	gctatcttcg	12480
gtttgttccc	actcacagga	gttaggagcg	cccttgggag	ctagagcggt	gtaggcagtg	12540
ccggagtagg	gcttaaaagt	aggccccctg	tcacgcacgc	cgcggtatgc	aaagtacgtg	12600
gaagccatat	caagcacacg	gttgtcacc	acagccaggg	tgaaccgcgc	tttgtacgag	12660
tacgcggtat	cctcgcggtc	cacagggatg	aaccgcagcg	tcaaacgctg	ggaccggtct	12720
gtggttacgt	cgtgcgtagg	tgcacccgtg	gggtttctaa	acttgttatt	caggctgaag	12780
tacgtctcgg	tggcgcgggc	aaactgcacc	agcccggggc	tcagggtact	cgaggcgtcc	12840
tggcccagaga	tgtgcatgta	agaccactgc	ggcatcatcg	aaggggtagc	catcttgga	12900
agcgggcgca	cggcggctca	gcagctcctc	tggcggcgac	atggacgat	acatgacaca	12960
tacgacacgt	tagctattta	gaagcatcgt	cggcgcttca	gggattgcac	ccccagaccc	13020
acgatgtctg	tcagtgtgct	ttgccagttg	ccactggcta	cgggcgcgat	cagctcgcca	13080
ccgctggcgg	cacggcgag	ggacgcgcgg	ctatggcggg	ttacaacaac	ggcggcggc	13140
cctggcagca	cagggtttctg	ctgggtgtca	gcggggggag	gcagggtccag	cgttacaggt	13200
gtgtgctggc	ccagcactcc	ggtagccatg	ggcgcgatgg	gacgggtggt	gggcaggcct	13260
tgtcttagtg	cctcctcgta	cgaggggagg	tcattctatt	gcgtcaccag	agtttcttcc	13320
ctgtcgggcc	gcggacgctt	ttcgccacgc	ccctctggag	acactgtctc	cacggccggt	13380
ggaggctcct	ctacgggagg	gcggggatca	agcttactgt	taactcttatt	ttgcactgcc	13440
tggttggcca	ggtccaccac	cccgctaatt	ccagaggcca	ggccatctac	caccttttgt	13500
tggaaaatttt	gctcttttcaa	cttgtccctc	agcatctggc	ctgtgctgct	gttccaggcc	13560
ttgctgccat	agttcttaat	ggtggaaccg	aaatttttaa	tgccgctcca	cagcgagccc	13620
cagctgaagg	cgccaccgct	catattgctg	gtgccgatat	cttgccagtt	ttccatgaac	13680
gggcgcgagc	cgtgtcgcgg	ggccagagac	gcaaagttga	tgtcttccat	tctacaaaaa	13740
agttacagga	ccaagcgagc	gtgagactcc	agacttttta	ttttgatttt	tccacatgca	13800
acttgttttt	aatcagtgtc	tctgcgcctg	caaggccacg	gatgcaattc	cgggcacggc	13860
gccaatcgcc	gcggcgatca	gtggaataag	gaggggcagg	ataccgcgcg	gcatgcgacg	13920
gtgcgacgcg	cgccgcgcgc	ggtggtgcgc	acgacgcgat	ccgcccgtca	ggccgtggcc	13980
ggccatgccc	ctcctacggt	gcattcttcc	tcggaatccc	ggcacccgga	aacggaggcg	14040
gcagggtgagg	gccatatctg	caagaaccac	aaagaccggc	ttttaaaacg	tgctgggggtg	14100
gtagcgcgct	gttggcagca	ccagggtcct	gcctccttcg	cgagccaccc	tgcgcacgga	14160
aatcggggcc	agcacgggct	ggcgacggcg	acggcggcgg	cgggttccag	tgggtggttcg	14220
gcgtcgggta	gtcgcctcgtc	ttctggggcg	gtaggtgtag	ccacgatagc	cgggggttagg	14280
cgcgatggaa	ggatgtaggg	catattcggg	cagtagtgcg	ctggcggtgc	cgtacttctc	14340
ggaacggcgc	gggcgcgggg	gggctgaaac	gcgaaacatc	cacgggtccg	tttgcacctc	14400
cgtagagggt	ttggacgcgg	ccgcagcggc	cgcctgcacc	gcggcatctg	ccaccgccga	14460
ggcaaccggg	gacgtttgtg	tctccatgcc	ctctgtggca	gtggcaatac	tagtgctact	14520

ggtggtgggt	atctgaacgt	ccacgggtctg	cacgcccagt	cccgggtgcca	cctgcttgat	14580
tggccgcacg	cggacctcgg	gctccagccc	aggctccacg	gtcattttttt	ccaagacatc	14640
ttccagtcgc	tggcgcttgg	gtaccatcag	ctgcacgggtg	ggtgccaaagt	caccagactc	14700
gcgcttttagg	ccgcgctttt	cttcgggacgg	tgcaagcgtg	ggcagcacct	gctgcagtgt	14760
cacgggcttt	aggctagggtg	ttgggttgcc	ctcgtccagc	ggcaacgcca	acatgtcctt	14820
atgccgcttt	ccgtaggcaa	actccccgag	gcgctcgttg	gcctgctcaa	gcaggtcctc	14880
gtcgcggtac	acctcatcat	acacgcgctt	gtagggtgcgg	gtggagcgct	caccgggctg	14940
aaaaactacg	gtggtgcccgg	gtcgcaaaac	acgtctttacg	cgtcgacctt	tccactgtac	15000
ccgcgcgctg	ggcgcggttg	cgtgcagcag	ttccacctcg	tcgtcaagtt	catcatcatc	15060
atcatctttt	tttttctttt	tgacctcgctt	tagcttttcgg	ggcttgtaat	cctgctcttc	15120
cttcttccggg	gggccataga	tctccggcgc	gatgacctgg	agcatctctt	ctttgatttt	15180
gcgcttgagg	atagcttcgt	tgcgcgcgcg	cgcgctggtg	tacatacaac	agtacgagtc	15240
taagtagttt	tttcttgcaa	tctagttgcg	cggggggcgg	gtgcgcacgg	gcacgcgcag	15300
gccgctaacc	gagtcgcgca	cccagtacac	gttgcccctg	cgaccttgag	tcatagcact	15360
aatggccgcg	gctgctgcgg	cggccgctcg	tcgcctggac	ctgggggggca	cagtgcacat	15420
acccgcggcc	agccttcgag	cggcccgcat	ggccgcccgt	cggccgggtgc	gacgtgcgcg	15480
gttaagcagg	gccgcgcggc	cgcgttgggc	ggcagtgccg	ggtcggcgag	gcggcgagac	15540
tgctacgcgc	ctccgcgcgc	tcttcatttt	agcataacgc	cgggctccgc	gcaccacggt	15600
ctgaatggcc	gcgtccactg	tggacactgg	tggcggcgctg	ggcgtgtagt	tgccgcgcctc	15660
ctccaccacc	gcgtcaatgg	cgatcatcgac	ggtggtgcgc	ccagtgcggc	cgcgcttggtg	15720
cgcgccccag	ggcgcgcggt	agtgcgccgcg	cacgcgcact	gggtgtttggt	cggagcgctt	15780
ctttgccccg	ccaaacatct	tgcttgggaa	gcgcaggccc	cagcctgtgt	tattgctggg	15840
cgatataagg	atggacatgt	ttgctcaaaa	agtgcggctc	gataggacgc	gcggcgagac	15900
tatgcccagg	gccttgtaaa	cgtaggggca	ggtgcggcgt	ctggcgctcag	taatggtcac	15960
tcgctggact	cctccgatgc	tgttgccgag	cggtagcgctc	ccgtgatctg	tgagagcagg	16020
aacgttttca	ctgacgggtg	tgatggtggg	ggctggcggg	cgcgccaaaa	tctggttctc	16080
gggaaaagcga	ttgaacacgt	gggtcagaga	ggtaaactgg	cggatgagct	gggagtagac	16140
ggcctggctg	ttgtagaagc	tcttgagtg	cacgggcaac	agctcggcgc	ccaccaccgg	16200
aaagttgctg	atctggctcg	tggagcggaa	ggtcacgggg	tcttgcatca	tgtctggcaa	16260
cgaccagtag	acctgctccg	agccgcaggt	tacgtcagga	gtgcaaagga	gggtccatga	16320
gcggatcccg	gtctgagggt	cgccgtagtt	gtatgcaagg	taccagctgc	ggtactgggt	16380
gaaggtgctg	tcattgctta	ttagggttgta	actgcgtttc	ttgctgtcct	ctgtcagggg	16440
tttgatcacc	ggtttcttct	gaggcttctc	gacctcgggt	tgccgagcgg	gggcggcagc	16500
ttctgcgcgt	gcctcggcct	cagcgcgctt	ctcctccgcc	cgtgtggcaa	aggtgtcgcc	16560
gcgaatggca	tgatcgttca	tgctctccac	cggctgcatt	gccgcggctg	ccgcgcttga	16620
gttctcttcc	gcgcgcgtgc	cactgttggt	gccgcgcctt	gcgccatccc	cgcctctgtt	16680
ggtgtcatct	tttaagcttg	cctggtaggc	gtccacatcc	aacagtgcgg	gaatgttacc	16740
accctccagg	tcatcgtagg	tgatcctaaa	gccctcctgg	aagggttgcc	gcttgccgat	16800
gcccaacaag	ttgctcagge	ggctgtgggt	gaagtccacc	cgcatcctg	gcagcaaaat	16860
gatgtctgga	tgggaaggctt	cgtttgatata	taccacaggc	atgacaagac	cagtgcactg	16920
gtcaaaacccc	agtctgaagt	tgcgggtgtc	aaactttacc	ccgatgtcgc	tttccagaac	16980
cccgttctgc	ctgccactt	tcaagtagtg	ctccacgatc	gcgttggttca	taagggtctat	17040
ggtcatggct	tcggagtagt	tgccctcggg	cagcgtgaac	tccacccact	catatttctag	17100
ctccacctgt	ttgtccttag	taagcgagcg	cgacaccatc	acccgcgcct	taaacttatt	17160
ggtaaacatg	aactcgttca	catttggtcat	ggtggtatgc	aggatgggtt	tcagggtcgcc	17220
gccccagtg	gaacggtcgt	caagattgat	ggtctgtgtg	cttgccctcc	ccgggctgta	17280
gtcattgttt	tgaatgaccg	tgggttagaaa	gttgctgtgg	tcgttctggt	agttccaggga	17340
tgccacatcc	gttgacttgt	tgtccacaag	gtacacacgg	gtggtgtcga	ataggggtgc	17400
caactcagag	taacggatgc	tgtttctccc	cccggtaggc	cgcagggtacc	gcggaggcac	17460
aaacggcggg	tccaggggag	catcgaaagg	ggaacccagc	gccgcgccca	ctggcgccgc	17520
gctcaccacg	ctctcgtagg	agggaggagg	accttcttca	tacatcgccg	cgcgctgcat	17580
actaagggga	atacaagaaa	accaacgctc	ggtgcccattg	ccttggtgag	tttttttattt	17640
tgcacatgc	tttttttttt	tttttttaaa	acattctccc	cagcctgggg	cgaagggtcg	17700
caaacggggt	gccactccct	cccaaatacca	ggacgctgct	gtcgtctgcc	gagtcacgt	17760
cctcccacac	cagaccccg	tgacggctgt	gcctttgacg	acgggtgggc	gggcgcgggc	17820
cgggcacatc	cctgtgctcc	tgcgcatacg	tcttccatct	actcatcttg	tccactaggc	17880
tctctatccc	gttggtggga	aatgccggag	gcaggttctt	ttcgcgctgc	ggctgcagca	17940
gcgagttggt	taggtactcc	tcttcgcccc	gcaggcgccg	gcggttggtg	cgagtgcctg	18000

taaaagaccc	tatcaagctt	ggaaatgggc	tactcgcctc	tgaccgcggg	gccgcagcgc	18060
ctagatcgga	caagctgctt	ggcctgcgga	agctttcctt	tcgcagcgcc	gcctctgcct	18120
gctcgcgctg	ttgcaactct	agcagggtct	gcgggtgcgg	ggaaaacacg	ctgtcgtcta	18180
tgtcgtccca	gaggaatcca	tcgttaccc	cgggcacctc	aaatcccccg	gtgtagaaac	18240
cagggggcgg	tagccagtgc	gggttcaaga	tggcatttgt	gaaatactcg	gggttcacgc	18300
cggccgcgcg	atgcaagtag	tccattaggg	gattgataaa	cggccggttt	gaggcataca	18360
tgcccggttc	catgttgccg	gcggtcattg	ccagcgccac	gctgggcgtt	accccgctcg	18420
gcatcagggt	aaggctcacg	ctctgctgca	catagcgcaa	gatgcgctcc	tcctcgctgt	18480
ttaaactgtg	caacgagggg	atcttctgcc	gccgggttgt	cagcaggtag	ttcagggttg	18540
cctccagggt	gcccgtgtcc	tcctgcccc	gcgcgcggct	gacacttgta	atctcctgga	18600
aagtatgctc	gtccacatgc	gcctgacct	tggcctcgcg	gtacagtgtc	agcaagtgc	18660
ctaggatatgt	gtcccgggac	acgctgccac	tgtccgtgaa	gggcgctatt	agcagcagca	18720
acaggcgcg	gttgggcgtc	agcaagctag	acacggtcgc	gcggtcgcct	gtgggagccc	18780
gcacccccca	cagcccctgc	aagttcttga	aagcctggct	cagggtttacg	gtctgcaggc	18840
cttgtctact	ggtctggaaa	aaatagtctg	gcccggactg	gtacacctca	ctttgcgggtg	18900
tctcagtcac	cattagccgc	agtgcgctca	caaagtttgt	gtagtctctc	tgtccccgcg	18960
gcacgttggt	ggcgtgtgta	ctcaggaagg	cgtttagtag	aaccatggag	ccagggttgc	19020
cctgctgctc	cgcgcgctca	cgctgcgcca	cggcctcgcg	cacatcccc	accagcgggt	19080
ccaggtttgt	ctgcacgttg	ccgctgttgt	aacgagccac	gcgctgaagc	agcgcgctcg	19140
agaccaggcc	ggcctcatcg	ggccggatgg	ccctgttttc	ggccagcgcg	tttacgatcg	19200
ccagcacctt	ctcgtgcgtg	gggtttgcgc	gcgcggggac	caccgcttcc	agaattgcgg	19260
agagccggtt	ggcctgcggc	tgctgcggga	acgcgtcagg	gttacgcgca	gtcagcgaca	19320
tgatgcgggtc	catgacctgg	cgccagtcgt	cgcgtggagt	aaggccggac	ggctggctct	19380
gcagcgcgcg	ccgcaccgcc	gggtccgttg	cgtcttgcac	catctgatca	gaaacatcac	19440
cgcttagtac	tcgccgtcct	ctggctcgta	ctcatcgctc	tcgtcatatt	cctccacgcc	19500
gccgacgttg	ccagcgcgcg	cgggtgccac	cgccagccca	ggtccggccc	cagctgcctc	19560
cagggcgcg	cggcttgggg	cccagcgcg	gtcagcgccc	gcgtcaaagt	aggactcggc	19620
ctctctatcg	ccgctgccc	tgccagccag	ggccctttgc	aggctgtgca	tcagctcgcg	19680
gtcgctgccc	tcgcgcggcc	ggctcacgct	gcgcgcttg	tggatgcgct	tggtgcgata	19740
aacgcccagg	tcgtcgctca	aggtaagcac	cttcaacgcc	atgcgcattg	agaacccttc	19800
gatctttacc	tccttgtcta	tgggaacgta	aggggtatgg	tatatcttgc	gggcgtaaaa	19860
cttgcccaga	ctgagcatgg	aatagttaat	ggcgccacc	ttgtcagcca	ggctcaagct	19920
gcgctcctgc	accactatgc	tctgcagaat	gtttatcaaa	tcgagcagcc	agcggccctc	19980
gggctctact	atgtttagca	gcgcacccct	gaatgcctcg	ttgtccctgc	tgtgtgcac	20040
tataaggaac	agctgcgcca	tgagcggctt	gctatttggt	ttttgctcca	gcgcgcttac	20100
aaagtccac	agatgcatca	gtcctatagc	cacctcctcg	cgcgccacaa	gcgtgcgcac	20160
gtggttgtta	aagctttttt	gaaagttaat	ctcctgggtc	accgtctgct	cgtaacgggt	20220
taccaggctg	gcggccgcca	cgtgtgcgcg	cgcgggacta	atcccggtcc	gcgcgtcggt	20280
ctcaaagtcc	tcctcgcgca	gcaaccgctc	gcggttcagg	ccatgccgca	actcgcgccc	20340
tgcgtggaac	tttcgatccc	gcactctcct	gggctcctct	ccctcgcggt	cgcgaaacag	20400
gttctgccc	ggcacgtacg	cctcgcgct	gtcagccttc	agctgcaccc	ttgggtgtcg	20460
ctcaggagag	ggcgctccta	gcgcgcccag	gccctcgccc	tcctccaagt	ccaggtagtg	20520
ccgggcccgg	cgccgcgggg	gttcgtaatc	accatctgcc	gccgcgtcag	ccgcggatgt	20580
tgccctcct	gacgcggtag	gagaagggga	gggtgccttg	catgtctgcc	gctgctcttg	20640
ctcttgccgc	tgctgaggag	gggggcgcac	ctgccgcagc	accggatgca	tctgggaaaa	20700
gcaaaaaagg	ggctcgctcc	tggttccgga	ggaatttgca	agcggggtct	tgcattgacg	20760
ggaggcaaac	ccccgttcgc	cgagtcggg	ccggcccag	actcgaaccg	ggggtcctgc	20820
gactcaaccc	ttggaaaata	accctccggc	tacaggggagc	gagccactta	atgctttcgc	20880
tttcacgcct	aaccgcttac	gccgcgcgcg	gccagtggcc	aaaaaagcta	gcgcagcagc	20940
cgccgcgcct	ggaaggaagc	caaaaggagc	gctccccctg	tgtctgacgt	cgcacacctg	21000
ggttcgacac	gcgggcccgt	accgcatgga	tcacggcgga	cggccggatc	cggggttcga	21060
accccggtcg	tccgccatga	tacccttgcg	aatttatcca	ccagaccacg	gaagagtgc	21120
cgcttacagg	ctctcctttt	gcaagggtcta	gagcgtcaac	gactgcgcac	gcctcaccgg	21180
ccagagcgct	ccgaccatgg	agcacttttt	gccgctgcgc	aacatctgga	accgcgtccg	21240
cgactttccg	cgcgccctca	ccaccgcgcg	cggcatcacc	tggatgtcca	ggtacatcta	21300
cggatatcat	cgcttatgt	tggaagacct	cgcgcccgga	gccccggcca	ccctacgctg	21360
gccccctctac	cgccagccgc	cgccgcactt	tttggtggga	tatcagtacc	tggtgcggac	21420
ttgcaacgac	tacgtctttg	actcaagggc	ttactcgctg	ctcagggtaca	ccgagctctc	21480

gcagccgggt	caccagaccg	ttaactgggtc	cgttatggcc	aactgcactt	acaccatcaa	21540
cacgggcgca	taccaccgct	ttgtggacat	ggatgacttc	cagtctaccc	tcacgcaggt	21600
gcagcaggcc	atattagccg	agcgcgttgt	cgccgacctg	gccctgcttc	agccgatgag	21660
gggcttcggg	gtcacacgca	tgggaggaag	agggcgccac	ctacggccaa	actccgccgc	21720
cgccgtagcg	atagatgcaa	gagatgcagg	acaagaggaa	ggagaagaag	aagtgcgggt	21780
agaaaggctc	atgcaagact	actacaaaga	cctgcgccga	tgtcaaaacg	aagcctgggg	21840
catggccgac	cgccctgcgca	ttcagcaggc	cggaccceaag	gacatgggtgc	ttctgtcgac	21900
catccgccgt	ctcaagaccg	cctacttttaa	ttacatcatc	agcagcacct	ccgccagaaa	21960
caaccccgac	cgccacccgc	tgccgcccgc	cacgggtgctc	agcctacctt	gcgactgtga	22020
ctgggttagac	gcctttctcg	agagggttttc	cgatccgggtc	gatgcggact	cgctcagggtc	22080
cctcggtggc	ggagtaccta	cacaacaatt	gttgagatgc	atcgttagcg	ccgtatccct	22140
gccgcacggc	agccccccgc	caaccataaa	ccgggacatg	acgggcggcg	tcttccaact	22200
gcgccccgc	gagaacggcc	gcgccgtcac	cgagaccatg	cgccgtcgcc	gcggggagat	22260
gatcgagcgc	tttgtcgacc	gcctcccgggt	gcgcccgtcgt	cgccgccgtg	ccccccctcc	22320
cccaccgccg	ccagaagaag	aagaagaagg	ggaggccctt	atggaagagg	agattgaaga	22380
agaagaggcc	cctgtagcct	ttgagcgcg	gggtgcgcgac	actgtcgccg	agctcatccg	22440
tcttctggag	gaggagttaa	ccgtgtcggc	cgccaactcc	cagtttttca	cggtatccgt	22500
ggacttctac	gaggccatgg	agcgccttga	ggccttgggg	gatatcaacg	aatccacgtt	22560
gcgacgctgg	gttatgtact	tcttctgtggc	agaacacacc	gccaccaccc	tcaactacct	22620
ctttcagcgc	ctgcgaaact	acgccgtctt	cgcccggcac	gtggagctca	atctcgcgca	22680
gggtggtcatg	cgcgcccgcg	atgccgaagg	gggcgtgggtc	tacagccgcg	tctggaacga	22740
gggaggcctc	aacgccttct	cgcagctcat	ggcccgcac	tccaacgacc	tcgcccgcac	22800
cgtggagcga	gccggacgcg	gagatctcca	ggaggagaag	atcgagcag	tcgagggcga	22860
aatcgccctat	caagacaact	caggagacgt	cgaggagatt	ttgcgccagg	cgccgctcaa	22920
cgacaccgaa	attgattctg	tcgaactctc	tttcagggttc	aagctcaccg	ggcccgtcgt	22980
cttcacgcag	aggcgccaga	ttcaggagat	caaccgccgc	gtcgtcgcg	tcgccagcaa	23040
cctccgcgcg	cagcaccagc	tcctgcccgc	gcgcggcgcc	gacgtgcccc	tgccccctct	23100
cccggcgggt	cccagcctcc	ccctacctcc	gggggcccgc	ccgcgtcacc	gcttttagat	23160
gcgcatccca	aggacacccc	cgcggcccac	gcgcccgccg	gcggtaaccg	agtcgcgcg	23220
cggggatgcg	gcctcttgca	agtcacgcac	gcccgcacca	accagcccc	ggaaatcagg	23280
tatcacctgg	acctagcccc	cgccctgacc	cggtctatgcg	aggtaaacct	gcaggagctc	23340
ccgcctgacc	tgtcgccgcg	ggagctccag	accatggaca	gctcccatct	gcgcgatgtt	23400
gtcatcaagc	tccgaccgce	gcgcgcggac	atctggactt	tgggctcgcg	cggcgtgggtg	23460
gtccgatcca	ccataactcc	cctcgagcag	ccagacggtc	aaggacaagc	agccgaagta	23520
gaagaccacc	agccaaaccc	gccaggcgag	gggctcaaat	tcccactctg	cttcttctgtg	23580
cgcggtcgtc	aggtaacct	cgtgcaggat	gtacagcccc	tgcaccgctg	ccagtactgc	23640
gcaggttttt	acaaaagcca	gcacgagtgt	tcggcccgtc	gcagggactt	ctactttcac	23700
cacatcaaca	gccactcctc	caactgggtg	cgggagatcc	agttcttccc	gatcggtctg	23760
catcctcgca	ccgagcgtct	ctttgtcacc	tacgatgtag	agacctatac	ttggatgggg	23820
gcctttggga	agcagctcgt	gcccttcatg	ctggttatga	agttcgccgg	agatgagcct	23880
ctggtgaccg	ccgcgcgaga	cctagccgtg	gaccttggat	gggaccgctg	ggaacaagac	23940
ccgcttacct	tctactgcat	caccccagaa	aaaatggcca	taggtcgcca	gttttaggacc	24000
tttcgcgacc	acctgcaaat	gctaattggc	cgtgacctgt	ggagctcatt	cgctcgcttcc	24060
aacctcatc	ttgcagactg	ggccctgtca	gaacacgggc	tcagctcccc	tgaggagctc	24120
acctacgagg	aacttaaaaa	attgcctctc	atcaagggca	ccccgcgctt	cttggaaactt	24180
tacatcgtag	gccacaacat	caacggcttc	gacgagatcg	tgctcgccgc	ccaggtaatt	24240
aacaaccgtt	ccgaggtgcc	gggaccttcc	cgcacacac	gcaactttat	gcctcgcgcg	24300
ggaaagatac	ttttcaacga	tgtcaccttc	gccctgccaa	acccgcgctt	caaaaagcgc	24360
acggactttt	tgtcttgga	gcagggcgga	tgcgacgaca	ctgacttcaa	ataccagtac	24420
ctcaaagtca	tggttaggga	cacctttgcg	ctcaccacac	cctcgctccg	gaaggccgcg	24480
caggcatatc	cgctaccggt	agaaaaggga	tgtctgcgct	accaggccgt	caaccagttc	24540
tacatgctag	gctcttaccg	ttcgagggcc	gacgggtttc	cgatccaaga	gtactggaaa	24600
gaccgcgaag	agtttgtcct	caaccgcgag	ctgtggaaaa	aaaagggaca	ggataagtat	24660
gacatcatca	aggaaacctt	ggactactgc	gccctagacg	tgcagggtcac	cgccgagctg	24720
gtcaacaagc	tgcgcgactc	ctacgcctcc	ttcgtgcgtg	acgcggtagg	tctcacagac	24780
gccagcttca	acgtcttcca	gcgtccaacc	atatcatcca	actcacatgc	catcttcagg	24840
cagatagtct	tccgagcaga	gcagccccgc	cgtagcaacc	tcgggtcccga	cctcctcgct	24900
ccctcgcacg	aactatacga	ttacgtgcgc	gccagcatcc	gcggtggaag	atgctaccct	24960

acatatcttg	gaatactcag	agagcccctc	tacgttttacg	acattttgcgg	catgtacgcc	25020
tccgcgcgtca	cccaccccat	gccatgggggt	ccccactca	accatacga	gcgcgcgcgtt	25080
gccgcccgcg	catggcagca	ggcgctagac	ttgcaaggat	gcaagataga	ctacttcgac	25140
gcgcgcctgc	tgcccgggggt	ctttaccgtg	gacgcagacc	ccccggacga	gacgcagcta	25200
gaccactac	cgccattctg	ttcgcgcaag	ggcgccgcgc	tctgctggac	caacgagcgc	25260
ctacgcggag	aggtagccac	cagcgttgac	cttgtaaccc	tgcacaaccg	cggttggcgc	25320
gtgcacctgg	tgcccagacga	gcgcaccacc	gtctttcccg	aatggcggtg	cgttgcgcgc	25380
gaatacgtgc	agctaaacat	cgcgcccaag	gagcgcgcgc	atcgcgacaa	aaaccaaacc	25440
ctgcgcctcca	tgcgcaagtt	gctgtccaac	gccctctacg	ggtcgtttgc	caccaagctt	25500
gacaacaaaa	agattgtctt	ttctgaccag	atggacgcgg	ccaccctcaa	aggcatcacc	25560
gcggggccagg	tgaatatcaa	atcctcctcg	tttttggaaa	ctgacaatct	tagcgcagaa	25620
gtcatgcccg	cttttgagag	ggagtactca	ccccaacagc	tggccctcgc	agacagcat	25680
gcggaagaga	gtgaggacga	acgcgcccc	accccccttt	atagcccccc	ttcaggaaca	25740
cccggtcacg	tggcctacac	ctataaacca	atcaccttcc	ttgatgccga	agagggcgac	25800
atgtgtcttc	acaccctgga	gcgagtggac	cccctagtgg	acaacgaccg	ctacccctcc	25860
cacttagcct	ccttcgtgct	ggcctggacg	cgagccttcg	tctcagagtg	gtccgagttt	25920
ctatacgagg	aggaccgcgg	aacaccgcgc	gaggacagcg	ctctcaagtc	gtatatacgg	25980
gacacggaca	gccttttctg	caccgagcgt	ggacacgggc	tcattggaac	cagaggttaag	26040
aaacgcataca	aaaagcatgg	gggaaacctg	gtttttgacc	ccgaacggcc	agagctcacc	26100
tggctcgtgg	aatgcgagac	cgtctgcggg	gcctgcggcg	cggatgccta	ctccccggaa	26160
tccgtatttc	tccgcgccc	gctctacgcc	cttaaaagtc	tgcactgccc	ctcgtgcggc	26220
gcctcctcca	agggcaagct	gcgcgccaag	ggccaacgcc	cggaggggct	ggactatgac	26280
accatgggtca	aatgctacct	ggccgacgcg	cagggcggaag	accggcagcg	cttcagcacc	26340
agcaggacca	gcctcaagcg	caccctggcc	agcgcgcagc	ccggagcgca	ccccctcacc	26400
gtgaccacaga	ctacgctgac	gaggaccctg	cgcccgctga	aagacatgac	cctggcccgct	26460
ctggacgagc	accgactact	gccgtacagc	gaaagccgcc	ccaaccgcgc	aaacgaggag	26520
atatgctgga	tccgagatgcc	gtagagcacg	tgaccgagct	gtgggaccgc	ctggaactgc	26580
ttgggtcaaac	gctcaaaaagc	atgcctacgg	cggacggcct	caaaccgctg	aaaaactttg	26640
cttccttgca	agaactgcta	tcgtctggcg	gcgagcgcc	tctggcgcat	ttggtcagg	26700
aaaacatgca	agtcaggggac	atgcttaacg	aagtggcccc	cctgctcagg	gatgacggca	26760
gctgcagctc	tcttaactac	cagttgcagc	cggtaatagg	tgtgatttac	gggcccaccg	26820
gctgcggtaa	gtcgcagctg	ctcaggaacc	tgctttcttc	ccagctgatc	tcccctaccc	26880
cggaaacgggt	tttcttcatc	gccccgcagg	tagacatgat	ccccccatct	gaactcaaag	26940
cgtgggaaat	gcaaactctgt	gagggttaact	acgcccctgg	gccggatgga	accattatac	27000
cttccttggt	caccctccgc	ccgcgctttg	taaaaatggc	ctatgacgat	ctactcctgg	27060
aacacaacta	tgacgttagt	gatcccagaa	atatcttcgc	ccaggccgcgc	gcccgtgggc	27120
ccattgccaat	cattatggac	gaatgcatgg	aaaatctcgc	aggtcacaag	ggcgtctcca	27180
agttcttcca	cgcatttctc	tctaagctac	atgacaaatt	tcccagtgct	accggatata	27240
ctgtgctgggt	ggttctgcac	aacatgaatc	ccgggagggg	tatggctggg	aacatagcca	27300
acctaataat	acagtccaag	atgcactctc	tatccccacg	tatgcacca	tcccagctta	27360
accgctttgt	aaacacttac	accaagggcc	tgcctctggc	aatcagcttg	ctactgaaag	27420
acatttttag	gcaccacgcc	cagcgtcctc	gtaacgactg	gatcatctac	aacaccaccc	27480
cgcagcatga	agctctgcag	tggtgctacc	tccaccccag	agacgggctt	atgcccatgt	27540
atctgaacat	ccagagtcac	ctttaccacg	tccctggaaa	aatacacagg	accctcaacg	27600
accgagaccg	ctggtcccgg	gcctaccgcg	cgcgcaaaac	ccctaataaa	agacagcaag	27660
acacttgctt	gatcaaaaatc	caaacagagt	ctgggttttt	tttatgtttt	aaaccgcatt	27720
gggaggggag	gaagccttca	gggcagaaac	ctgctggcgc	agatccaaca	gctgctgaga	27780
aacgacatta	agttcccggg	tcaaagaatc	caattgtgcc	aaaagagccg	tcaacttgct	27840
atcgcggggc	gatgaacggg	aagctgcact	gcttgcaagc	gggctcagga	aagcaaagtc	27900
agtcacaatc	ccgcgggcgg	tggtgcagc	ggctgaagcg	gcgcgggagg	ctgcagctc	27960
caacggcggt	ccagacacgg	tctcgtaggt	caaggtagta	gagtttgccg	gcaggacggg	28020
gcgaccatca	atgctggagc	ccatcacatt	ctgacgcacc	ccggcccatg	ggggcatgcg	28080
cgttgtcaaa	tatgagctca	caatgcttcc	atcaaacgag	ttggcgctca	tggcgggcgc	28140
tgctgcaaaa	cagatacaaa	actacatgag	acccccacct	tatatattct	ttccaccctc	28200
taagccccgc	ccatcgatgg	caaacagcta	ttatgggtat	tatgggtgct	agcgacatga	28260
ggttgccccg	tattcagtg	cgctgatttg	tattgtctga	agttgttttt	acgttaagtt	28320
gatgcagatc	aattaatacg	atacctgcgt	cataattgat	tatttgacgt	ggtttgatgg	28380
cctccacgca	cgttgtgata	tgtagatgat	aatcattatc	actttacggg	tcctttccgg	28440

```

tgatccgaca ggttacgggg cgccgacctc gggggttttc gctatttatg aaaatatttc 28500
ggtttaaggc gtttccgttc ttcttcgtca taacttaatg tttttattta aaataccctc 28560
tgaaaagaaa ggaaacgaca ggtgctgaaa gcgaggcttt ttggcctctg tcgtttcctt 28620
tctctgtttt tgtccgtgga atgaacaatg gaagttaacg gatccaggcc gcgagcaaaa 28680
ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc 28740
cgccccctg acgagcatca caaaaatcaa cgctcaagtc agagggtggcg aaaccgaca 28800
ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg 28860
accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct 28920
catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt 28980
gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtacta tcgtcttgag 29040
tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc 29100
agagcgaggc atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac 29160
actagaagaa cagtatttgg tatctgcgct ctgccaaagc cagttacctt cggaaaaaga 29220
gttggttagct cttgatccgg caaacaacc accgctggta gcggtggttt ttttgtttgc 29280
aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg 29340
gggtctgacg ctcagtggaa cgaaaactca cgtaaaggga ttttggtcat cagattatca 29400
aaaaggatct tcacctagat ctttttaaata taaaaatgaa gttttaaatc aatctaaagt 29460
atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtaggc acctatctca 29520
gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtagtgtg gataactacg 29580
atacgggagg gcttaccatc cgccccagt gctgcaatga taccgcgtga cccacgctca 29640
ccggtccttg atttatcagc aataaaccag ccagccggaa gtgccgagcg cagaagtggg 29700
cctgcaactt tatccgcctc catccagtct attagtgtgt gccgggaagc tagagtaagt 29760
agttcgccag ttaatagttt tcgcaacggt gttgccattg ctacaggcat cgtgggtgtc 29820
cgctcgctcg ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca 29880
tgatcccca tggtgtgcaa aaaagcgggt agctccttcg gtccctccgat agttgtcaga 29940
agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact 30000
gtcatgccat ccgtaagatg cttttctgtg actggtgagt attcaaccaa gaatacggga 30060
taataccgcg ccacatagca gaactttaaa agtgctcatc attgggaaac gttcttcggg 30120
gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgcgc 30180
accaagtgat tcttctgcat cttttacttt caccagcgtt tctgggtgag caaaaacagg 30240
aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact 30300
tttccttttt caatattatt gaagcattta tcagggttat tgtctcatca gcggatacat 30360
atttg                                     30365

```

```

<210> 5
<211> 33
<212> DNA
<213> Homo sapiens

```

```

<400> 5
gcggaattcg gcttggtgac ttagagaaca gag 33

```

```

<210> 6
<211> 33
<212> DNA
<213> Homo sapiens

```

```

<400> 6
gcgggatcct tgaaccgga ccctctcaca cta 33

```

```

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

```

```

<220>

```

<223> derived from Adenovirus

<400> 7

actctcttcc gcatcgctgt

20

<210> 8

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 8

cttgcgactg tgactgggta g

21

<210> 9

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 9

ccgcacccac tatcttcata

20

<210> 10

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 10

ggtgtccaaa gggtcggaga

20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/27682

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :C12N 15/86, 15/861, 5/10, 15/11, 15/63, 15/64, 15/65; A61K 48/00

US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/93.1, 93.2, 93.6; 435/320.1, 69.1, 455, 456, 457, 325, 369, 91.1, 91.4, 91.42

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	US 6,080,576 A (ZAMBROWICZ et al.) 27 June 2000 (27.06.00), see entire document, especially Figure 2, claims 1-3 and columns 10 and 16.	1-2, 4-11, 13
A	US 5,919,676 A (GRAHAM et al.) 06 July 1999 (06.07.99), see entire document, particularly claims 1-6 and columns 2 and 5-6.	1-34
A	AGAH et al. Gene recombination in postmitotic cells. Journal of Clinical Investigation. July 1997, Vol. 100, No. 1, pages 169-179, especially pages 171-173.	1-34

☐

Further documents are listed in the continuation of Box C.

☐

See patent family annex.

"A"	document defining the general state of the art which is not considered to be of particular relevance	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search

01 NOVEMBER 2001

Date of mailing of the international search report

31 DEC 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

David Guzo

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/27682

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

424/93.1, 93.2, 93.6; 435/320.1, 69.1, 455, 456, 457, 325, 369, 91.1, 91.4, 91.42

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST, Dialog, NTIS, Medline, Biotech, Biosis, Biosci, Chemical Abstracts

Search terms: adenovirus, recombinase, target site, Cre, cleavage site, gene therapy